DOCUMENT RESUME

ED 162 279 CS 004 483

AUTHOR	Klopp, Pamela M.; And Others
TITLE	Technical Manual for Forms P and Q of the Wisconsin Tests of Reading Skill Development: Comprehension. Technical Report No. 465.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Individualized Schooling.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Sep 78
CONTRACT	NIE-G-78-0117
NOTE	130p.; Some tables may not reproduce well due to small type
EDRS PRICE	MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS	*Criterion Referenced Tests; Educational Assessment; Elementary Education; Measurement Techniques; Readability; *Reading Comprehension; *Reading Development; *Reading Skills; *Test Construction
IDENTIFIERS	*Wisconsin Tests of Reading Skill Development

ABSTRACT

The procedures used in developing test forms P and Q of the "Wisconsin Tests of Reading Skill Development: Comprehension" are described in this report. The criterion-referenced tests assess the 40 skills and objectives of the "Wisconsin Design for Reading Skill Development: Comprehension" commercial edition. The report contains individual test descriptions as well as general descriptive information regarding readability specifications, item content and test format, test directions, scoring and use of test results, and the tryout population. A rationale for test criteria, estimated summary statistics for both test forms, and actual summary statistics determined by administering the commercial edition are provided. Appendixes include the statement of skills and objectives for the test, estimated summary statistics for forms P and Q, item writing guidelines, and a list of selected "Design"-related studies. (Author/FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED162279

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report No. 465

TECHNICAL MANUAL FOR FORMS P AND Q OF THE WISCONSIN TESTS OF READING SKILL DEVELOPMENT: COMPREHENSION

by

Pamela M. Klopp, Craig L. Hougum,
Deborah M. Stewart, and Karlyn Kamm

Report from the Project on
Studies in Language: Reading and Communication

Wayne Otto
Faculty Associate

CS004483

Wisconsin Research and Development
Center for Individualized Schooling
The University of Wisconsin
Madison, Wisconsin

September 1978

Sections I, II, and III are from the preliminary technical manual by Pamela M. Klopp (1977). Section IV was adapted by Karlyn Kamm from the Word Attack and Study Skills technical manual (Stewart, 1976). Data analysis for Section V was carried out by Craig L. Hougum.

© 1978 - The Regents of the University of Wisconsin System for the Wisconsin Research and Development Center for Individualized Schooling.
Copyright is claimed only during the period of development, test, and evaluation, unless authorization is received from the National Institute of Education to claim copyright on the final materials. For the current copyright status, contact either the copyright proprietor or the National Institute of Education.

Published by the Wisconsin Research and Development Center for Individualized Schooling, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. OB-NIE-G-78-0117

MISSION STATEMENT

The mission of the Wisconsin Research and Development Center is to improve the quality of education by addressing the full range of issues and problems related to individualized schooling. Teaching, learning, and the problems of individualization are given concurrent attention in the Center's efforts to discover processes and develop strategies and materials for use in the schools. The Center pursues its mission by

- conducting and synthesizing research to clarify the processes of school-age children's learning and development
- conducting and synthesizing research to clarify effective approaches to teaching students basic skills and concepts
- developing and demonstrating improved instructional strategies, processes, and materials for students, teachers, and school administrators
- providing assistance to educators which helps transfer the outcomes of research and development to improved practice in local schools and teacher education institutions

The Wisconsin Research and Development Center is supported with funds from the National Institute of Education and the University of Wisconsin.

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR INDIVIDUALIZED SCHOOLING

TABLE OF CONTENTS

		Page
	List of Tables	vii
	List of Figures	ix
	Abstract	xi
I.	Background	1
II.	R & D Center Test Development Procedures	5
	Multiple-Choice Tests	5
	Completion Tests	8
III.	Descriptive Characteristics of the Assessment Materials	11
	Readability Specifications	11
	Item Content and Test Format	11
	Directions for Administering the Tests	13
	Scoring and Use of Test Results	15
	Individual Test Descriptions	16
IV.	Rationale for Test Criteria	35
	Content Specifications	35
	Usability Specifications	37
	Psychometric Characteristics	38
	Summary	41
V.	Summary Statistics for the 1977 Commercial Edition of WTRSD Comprehension	43
	References	93

TABLE OF CONTENTS

Page

Appendices

A. Statement of Skills and Objectives for the Wisconsin Design for Reading Skill Development: Comprehension Commercial Edition 95

B. Estimated Summary Statistics for Forms P & Q 103

C. Item Writing Guidelines for the Wisconsin Tests of Reading Skill Development: Comprehension Commercial Edition 117

D. List of Selected Design-related Studies. 123

LIST OF TABLES

Table		Page
1	Approximate Testing Times for WTRSD: Comprehension Battery, Levels A-G	14
2	Appropriate Performance Categories and Types of Instruction	16
Summary Statistics for 1977 Commercial Edition of WTRSD: Comprehension		43
3-9	Test Statistics (Mean, Standard Deviation, Hoyt r, Standard Error of Measurement, and Intercorrelations, Levels A-G	44-50
10-16	Frequency Distributions of Raw Scores and Number and Percent Mastering from None to All Skills, Levels A-G	51-57
17-23	Frequency Distrubutions of Proportion Correct and Proportion Omitting Each Item, Levels A-G	58-64
24-30	Frequency Distributions of Item-Test Correlations, Levels A-G	65-71
31-37	Parallel Forms Mastery-Decision Characteristics, Levels A-G	72-78
38-44	Sample Characteristics for Parallel Forms Analysis, Levels A-G	79-85
45-51	Parallel Forms Conditional Probabilities for Mastery, Levels A-G	86-92

LIST OF FIGURES

Figure		Page
1	WDRSD: Comprehension Strands, Commercial edition	3
2	R & D Center Model for Development of Criterion-Referenced Tests	5

I

BACKGROUND

The Wisconsin Design for Reading Skill Development (Design) is an objective-based system that can provide both structure and substance for an elementary school reading program, kindergarten through grade 6. The Design provides the basis for a skill-oriented approach to the teaching of reading and is based on the assumption that if learners acquire and apply the essential subskills, they will be successful readers. The six Design skill areas are Word Attack, Study Skills, Comprehension, Self-directed Reading, Interpretive Reading, and Creative Reading. The Design's four major components are (1) skills and objectives; (2) assessment materials; (3) instructional resources; and (4) management techniques and materials. A complete description of the Design, a rationale for its development, and guidelines for its implementation are provided in the Rationale and Guidelines (Otto & Askov, 1974).

The Wisconsin Tests of Reading Skill Development (WTRSD) are the Design's formal assessment component. The WTRSD were developed to assess the behavioral, or prescriptive, objectives in Word Attack, Comprehension, and Study Skills and, therefore, are criterion-referenced. Each test measures a single behavioral objective. The majority of the WTRSD are paper-and-pencil tests which can be group-administered and scored either by hand or by computer. There are two parallel test forms, P and Q, for most of the objectives. The tests are available either in booklet format (for all the objectives at a given skill level) or in single-sheet format (for a single objective).

The booklets are used for break-in testing to find the students' initial instructional levels, and the single-sheet tests are used after skill instruction to assess individual attainment of criterion performance. Manuals for test administrators are included with the assessment materials.

Test forms P and Q of WTRSD: Comprehension assess the 40 skills and objectives of the WDRSD: Comprehension commercial edition. The tests are a second major edition and represent a substantial redevelopment effort. Like Word Attack and Study Skills, the tests are assigned to levels which correspond roughly to traditional elementary grade levels: Level A indicates kindergarten; Level B, grade 1; Level C, grade 2; Level D, grade 3; Level E, grade 4; Level F, grade 5; and Level G, grade 6.

The commerical edition of WDRSD: Comprehension is a refinement of a developmental edition that underwent a pilot test during 1972-73, a small-scale field test during 1973-74, and a large-scale field test during 1974-76. Sufficient evaluative information about the skills, objectives, and WTRSD had been collected by the end of the small-scale field test to suggest a need for redevelopment, even though the developmental edition in use during 1974-76 was a respectable version. The redevelopment effort was guided by a comprehensive review of the professional literature; a careful examination of standardized tests, basal readers, and content materials; research studies; expert outside review; and field experiences with the developmental edition.

The strands and substrands of the 40 skills that emerged from the redevelopment are shown in Figure 1.

Main Strands	Substrands	Levels of Skills and Objectives
Word Meaning	Word Parts	- - - - E F G
	Context Clues	- - - D E F G
Sentence Meaning	Detail	- B C D E - -
	Paraphrase	- - C D E F G
Passage Meaning	Central Thought	A B C D E F G
	Relationships/ Conclusions	A B C D E F G
Sequence		A B C D E F G

Figure 1. WDRSD: Comprehension strands, commercial edition.

The behavioral objective for each skill is listed by level in Appendix A. A discussion of how the commercial edition evolved from the developmental edition and a rationale statement for each of the strands in Figure 1 is given in Chapter 2 of the Teacher's Planning Guide: Comprehension (Otto & Kamm, 1977).

The skills and objectives of the commercial edition retain two general characteristics of the developmental edition: They are limited to literal listening (Levels A and B) and reading comprehension (Levels C-G), and they require convergent thinking. The sequential arrangement of skills within strands for the commercial edition was established from logical analysis, the developers' judgment, and tryout data. The commercial edition of Comprehension differs from the Word Attack and Study Skills elements in that the skill mastery level of 80 percent correct responses has been replaced with performance bands. A learner who performs at 80 percent or above on a criterion test is a skill Applier; a learner who performs between 50-79 percent is a Pre-Applier; and a learner who performs below 50 percent is a Novice. The implementation strategies for the commercial edition (Chapter 3, Otto & Kamm, 1977) emphasize not only skill development but also application, utilizing a variety of reading and content area materials of varying levels of difficulty.

II

R & D CENTER TEST DEVELOPMENT PROCEDURES

Multiple-choice Tests

Because the Design is primarily a management system, its criterion-referenced tests must be of the highest quality. Development, tryout, and evaluation of WTRSD: Comprehension was a highly specialized activity that approximated the Center model for the development of criterion-referenced tests shown in Figure 2.

STEP	ACTIVITY
1	Development of test specifications and sample items.
2	Writing of sufficient items to yield one test form.
3	Pilot administration of items from step 2 to a small group of learners.
4	Formal administration of revised items to a number of learners large enough to gather item data.
5	Assembly of a test administrator's manual.
6	Test use under normal conditions by a number of schools or classes.

Figure 2. R & D Center model for development of criterion-referenced tests.

Since the tests are criterion-referenced, test specifications followed from the Comprehension behavioral objectives with which the tests were to be used. Parallel forms, Forms P and Q, were developed for the multiple-choice tests. Therefore, an additional item-writing step occurred after step 3 to ensure an adequate pool of items from which Forms P and Q could be assembled. Data presented in Section V were gathered during battery testing.

Each of steps 1-5 in Figure 2 included a review phase in addition to the developmental phase. The internal Center review phase was guided by these evaluation questions: (1) Do the tests reflect the behavioral objectives? (2) Are item formats appropriate? (3) Is the intent of the items clear? (4) Are the tests appropriate to their audience? (5) Are the test statistics acceptable? (6) Are the test directions clear? and (7) Is the testing time appropriate? In addition to the internal Center review, an external review of test items was conducted by an urban reviewer from Milwaukee, Wisconsin, prior to undertaking step 4. The primary purpose of the urban review was to identify items that would stymie a city child because the content was unfamiliar. A review of test directions, to determine whether the task would be clear to learners across a range of grade (reading) levels in which the test might be administered, was a secondary focus of the urban review.

Pilot tryouts (step 3) to small groups of average learners at the grade level for which the skills-objectives were intended supplemented the review phase by providing early indicators of the appropriateness of item formats, clarity of items and test directions, difficulty level, etc. Large-scale tryouts (step 4) with 100 or more learners of the target popula-

tion provided several kinds of statistical data for evaluating each test: obtained values for mean, standard deviation, internal consistency reliability, standard error of measurement, and number and percent of learners performing at 80 percent or above. A reliability estimate of .85 or above on the tryout edition ensured a satisfactory reliability estimate for the final test forms. A test was considered to be appropriately leveled if 30 percent or more of the tryout population achieved a score of 80 percent or above on the tryout test. The number and proportion of testees selecting each choice, and biserial and point biserial correlations with the total test score, were obtained for the answer choices on individual items within each test. Biserial correlations equal to or greater than .50 for correct answer choices and equal to or less than -.20 for incorrect answer choices were considered acceptable item statistics.

Feedback from teachers was an additional source of tryout information. School personnel who administered the tests were asked routinely to report the amount of time required for administration and were invited to comment on specific test items and test directions. For example, they were encouraged to note which vocabulary words the students wanted pronounced most often and whether any of the pictures at Levels A and P were confusing.

Step 5 in the test development cycle was accomplished by selecting from the total set of items administered in step 4 sufficient items to make parallel forms. Content specifications in addition to acceptable item statistics guided the item selection procedure. The items selected were then arranged into Forms P and Q. For these items, the proportion of testees selecting the correct choice and the point biserial $\underline{r}$ for the

correct choice were used to obtain split-half estimates for the mean, standard deviation, and reliability of the two forms. The technique by which the tests were divided into parallel forms is described in An Index of Parallel Forms of Criterion-Referenced Tests as a Function of the Distribution of Test Scores (Junker & Stewart, 1973). The procedure for rescoring the total test to obtain estimates of inter-form reliability and the proportion agreement between forms was not carried out.

The estimated summary statistics for Forms P and Q of the multiple-choice tests are given in Appendix B. These estimates were prepared after steps 1-5 had been completed. Statistics gathered during normal battery testing are given in Section V.

Completion Tests

The test development sequence outlined thus far was followed for each of the 33 multiple-choice Comprehension tests. Program usability is enhanced by the availability of multiple-choice tests for skill assessment since they can be quickly and easily scored. A multiple-choice format was not feasible, however, for assessing the other seven skills. Therefore, a group-administerable, written completion format was used for each of six tests: E1-Word Parts--Prefixes; F1-Word Parts--Suffixes; G1-Word Parts--Combining Forms; G2-Word Parts--Word Roots; G4-Paraphrase: Restatements; and G5-Central Thought: Generation. An individually administered, oral completion assessment exercise was developed for G7-Conclusions: Modification.

The test administrator must use subjective judgment when assessing a learner's response for each of these seven tests. Therefore, the collection of usability information about test scoring was a primary concern. It was important, for example, to learn whether the scoring guidelines were clear and adequate. Teachers were asked not only to administer the tests but also to score them, using guidelines for scoring provided by the devel-

opers. The tests were then returned to the Center, and the teacher-scored results were withheld until the developer had also scored the tests. Then the scoring guidelines were evaluated by examining (1) the degree of agreement between teacher-scoring and developer-scoring of individual test items, (2) teacher-developer agreement on the total raw score for each student, and (3) whether the student should be designated an Applier (total percent correct of 80 or above). In addition, teachers were asked to provide testing time and scoring time estimates, answer specific questions about test administration and scoring, and make recommendations, if necessary, to improve the test directions and scoring information. Thus, the degree of teacher involvement in these tryouts was greater than is normally required.

The development of these seven tests differed from the model in two other ways. First, there was only one tryout administration, rather than a pilot administration followed by a formal administration to a large number of learners. The testees were teacher-selected on the basis of "average" reading ability, or "at" grade level proficiency. Second, except for Test G7-Conclusions: Modification, parallel forms in terms of content specifications were developed prior to tryout instead of assembled from one tryout edition. No learner was administered both forms of a test.

Learner responses to items in Tests G2, G4, and G7 are given either full credit or no credit. Partial credit is permitted for the G5 test items and for parts of Test E1, F1, and G1. Data analysis involved computing (1) raw score total and mean score for individual test items across the total number of testees and (2) total raw score and percent correct score for each testee. Each data analysis step was conducted for both the teacher-scored results and the developer-scored results. For each test the proximity of the total raw score means for the two forms was carefully evaluated.

III

DESCRIPTIVE CHARACTERISTICS OF THE ASSESSMENT MATERIALS

Readability Specifications

In order to provide guidelines for item writing, readability specifications were developed with respect to (1) general sentence structures, (2) phrase limitations, (3) sentence length, and (4) vocabulary restrictions. The sentence structures that were permitted increase in complexity at each successive skill level, as do the maximum and average sentence lengths and maximum words per passage. Vocabulary restrictions which were translated into four categories of words, pertained to word lists and decoding level. Words which fell into three of the categories were considered "familiar"; those which fell into the fourth were considered "unfamiliar." The item writing guidelines for each of the four basic aspects are detailed in **Appendix** C.

Item Content and Test Format

The items are composed mainly of developer-written narrative based on experiences common to elementary-age learners. The purpose of the items is to assess literal Comprehension skills, not the knowledge learned from a curriculum. In addition to narrative, tests in the Central Thought and Relationships/Conclusions substrands incorporate science and social studies content either written by the developer or taken from commercially published textbooks. Again, however, the content serves as a vehicle for the assessment of skills, not knowledge.

The framework for assessment mirrors the developmental organization of each strand. For example, the learner answers questions based on passages in the Central Thought and Relationships/Conclusions tests of the Passage Meaning strand. In the Central Thought tests, passage organization--presence or absence of (1) a topic sentence or central thought statement, and (2) relevant or irrelevant details--varies from level to level depending on the skill being assessed. In the Relationships/Conclusions tests, the kinds of relationships given--class inclusive, comparison-contrast, conditional, cause-effect--and whether these relationships are directly or indirectly stated, are also variable. Similarly, sentence structure is the vehicle for assessing the Detail and Paraphrase skills in the Sentence Meaning strand, with the structures prescribed by the objectives becoming more complex at each higher level.

The test format changes from pictures to words to accommodate the learner's level of decoding ability. For example, the Level A tests, which are intended for children at the reading readiness stage, do not require word reading. Instead, the child listens to a short story read by the test administrator, and then selects the answer from a row of three pictures. Each row/item is marked with a small place-holder picture, instead of a numeral, to help the child keep his or her place. At Level B, three tests have picture formats and two tests have written-oral formats. In a written-oral format the items printed on a test page are read silently by the child as the test administrator reads them out loud. Students at Levels C-G are required to read material independently since their decoding ability is more advanced and automatic.

Response options, example items, and test sequence are other characteristics. The child selects from multiple options for 33 tests, generates written responses ranging from word parts to complete sentences for six tests, and responds to questions orally for one test. Answers are marked directly on the test pages at Levels A-D. Reusable booklets and separate answer sheets are used at Levels E, F, and G. Every test has at least one example item that the testees work through together with the test administrator, to gain experience in marking answers in addition to becoming oriented to the task. The test sequence within a skill level simply follows the organization of the strands. That is, the test order within a level is not based on inherent or empirically derived difficulty sequences.

Directions for Administering the Tests

A separate test administrator's manual for each level provides information about prescribed conditions, materials and preparations needed, directions for administering Form P and Form Q of each test, and scoring.

Since the example items specific to each test are identifical for Forms P and Q, the same general directions are used to administer both forms. The general directions, which are read out loud, state what the test is about, explain the example items fully, and tell the testee how to proceed. The specific story or item directions for the orally administered tests at Levels A and B follow the general directions by test form. Although the student must be able to read the test items independently at Levels C-G, the test administrator is permitted to help with word pronunciation.

The average time required for completion of a test is approximately 15-20 minutes. The testing time for each test was estimated from teachers' indications of the actual time spent administering the tryout edition. Since the tests are not timed and children are encouraged to take time to do their best, this information is presented for planning purposes only. Table 1 provides approximate testing time for administration of the complete battery at each level.

TABLE 1

APPROXIMATE TESTING TIMES FOR WTRSD: COMPREHENSION BATTERY, LEVELS A-G

Test Level	No. of Tests	Group Size[a]	No. of Sittings	Approximate Total Testing Time (Minutes)
A	3	10	2	35
B	5	15	3	75
C	5	20	3	75
D	6	25	3	90
E	7	25	4	140
F	6	25	3	125
G	7	25	4	145[b]

[a] Recommended maximum number of children to be tested by a single test administrator.

[b] Does not include the 20 minutes estimated for Test G7-Conclusions: Modification, which must be administered individually to each testee.

The complete battery of either Form P or Form Q is used to determine each learner's initial needs for instruction in Comprehension. It is also used to assess performance at successive levels as the learner "completes" (develops and applies) the skills. Each sitting for battery testing is typically composed of two or three tests.

Scoring and Use of Test Results

Scoring keys are identical for Forms P and Q of the multiple-choice tests. The scorer's subjective judgment is necessary in assessing the responses for tests E1, F1, G1, G2, G4, G5, and G7. Guidelines are provided to assist the scorer in making the necessary judgments. These guidelines include a sample response for each item of each test form. Information about basic criteria, grammar, content, and recording scores is also given for these seven tests.

Test example items are never scored. A conversion chart that accompanies the scoring information in each manual is used to determine the percentage of items correct on each test. Results from battery testing (preassessment) are recorded on the Comprehension Pupil Profile Card and then translated into the performance bands shown in Table 2.

Each learner's results from battery testing are used to record performance categories, verify his or her performance group, determine whether retesting is needed, and determine appropriate instruction (Chapter 3, Otto & Kamm, 1977).

Test results are also used to assess the performance levels of Novices and Pre-Appliers after instruction. The test form not used for battery testing is recommended for use in posttesting.

TABLE 2

APPROPRIATE PERFORMANCE CATEGORIES* AND TYPES OF INSTRUCTION

Percentage Score	Performance Category	Type of Instruction Most Appropriate
0-49	Novice (N)	Developmental work, teacher's resource file activities
50-79	Pre-Applier (PA)	Combination of developmental activities and application activities which are teacher-directed or closely supervised
80-100	Applier (A)	Independent application activities carefully specified by the teacher

*Indicates a student's level of performance with regard to a skill.

(Otto, W. & Kamm, K. Wisconsin Design for Reading Skill Development: Teacher's Planning Guide--Comprehension. Minneapolis: National Computer Systems, 1977.)

Individual Test Descriptions

Specific tests at certain levels share similar characteristics which are described here and are not repeated in the individual descriptions. The content of the six picture tests--A1, A2, A3, B3, B4, and B5--includes illustrations of named people, scenes, and objects which are clearly recognizable and contain a minimum of detail. An effort was made to provide a balance with respect to gender and ethnicity in the test items depicting people. The stimulus sentences or stories for the picture items are not printed on the test page but are read aloud by the test administra-

tor. Tests B1 and B2 are written-oral; that is, children follow the written copy on their test page as it is read out loud by the test administrator. As mentioned previously, all test copy at Levels C-G is read by students independently, although the test administrator may help with word pronunciation if necessary.

Individual test descriptions are presented in skill level and number sequence.

Test A1--Central Thought--Topic: Pictures. The child selects the one of three pictures that best illustrates the topic of a four- or five-sentence passage that is read to him or her. There are two types of passages, dependent (five per test form) and independent (seven per test form).[1] Each of the 12 passages has an organizer (topic sentence) which is the first sentence in 11 passages and the last sentence in one passage. Incorrect answer choices depict details from the story. The test administrator reads each passage once.

Test A2--Relationships. The child selects the one of three pictures that depicts a synthesis of the information given in a statement of relationship read by the test administrator. Four items assess the class-inclusive relationship--one that characterizes an object or class by the words every, all, each, and none. The remaining 11 items assess the comparison-contrast

---

[1] In a "dependent" passage, the topic can be identified only by synthesizing all of the relevant details. For example, in a passage that describes a process or action, such as "cleaning the house," no relevant detail by itself conveys the topic; the topic is conveyed only by synthesizing all of the relevant details. In an "independent" passage, the topic can be identified by the information given in each relevant detail. For example, in a passage that describes an object, such as "a wagon," each relevant detail by itself conveys the topic in some way.

relationship--one of degree or equality--as characterized by the words smaller, wider, bigger, narrower, same, longer, shorter, higher, lower, more, and less. There is one item for each relationship word listed. Incorrect answers for the class-inclusive items are pictures in which the relationship is completely absent or only partially correct. For the sentence, "Each of Kim's sneakers has a hole in it," for example, pictures of no sneakers with holes and some sneakers with holes would be incorrect answers. Incorrect answers for the comparison-contrast relationships depict either a reversal of the comparison being questioned or an equivalence so that comparison-contrast is not possible. Each statement of relationship is read twice.

Test A3--Sequence--First/Last Event. After listening to a story with three simple events, the child selects the one of three pictured story events that shows what happened either first or last. The answer choices cannot be ordered without the accompanying story. Connectors "and" or "and then" combine the three independent events into the one-sentence story. Approximately half of the 15 story questions ask what happened first and half ask what happened last. The item order for these two question types was randomly assigned. Picture order is controlled so that the first and last event pictures occur nearly equally in the first, second, and third position in the rows of pictures. The test administrator reads each story twice and each question once.

Test B1--Detail. The child answers 15 questions about specific details in five stories. The stories are written in simple sentences that are in the active voice; the questions do not paraphrase the story content. The parts of speech tested include objects of positional prepositions, modifiers, direct objects, and verbs. Samples of the same part of speech being ques-

tioned are provided as the two or three incorrect answers for each item. The samples are either taken directly from the stories or represent a logical completion to the question stem. Stories, questions, and answer choices are read by the test administrator only once.

Test B2--Central Thought--Topic: With Organizer. The child identifies the topic of each of 15 passages. There are two types of passages, dependent (4 per test form) and independent (11 per test form). (See footnote 1 for a definition of dependent and independent passages.) Each passage has an organizer (topic sentence) which is the first, middle, or last sentence in approximately one-third of the passages. The topic, selected from three or four answer choices, is expressed in a single word or a two- to four-word phrase. Incorrect choices are one- to three-word phrases about the details in the passage. The test administrator reads the passages and answer choices once.

Test B3--Relationships/Conclusions: Outcomes. The child selects the one of four pictures that shows the action likely to follow a three-part event described in a sentence read by the test administrator. The correct answer shows all parts of the event and the outcome. Wrong answer choices depict nonsense outcomes, part of the event missing, or contradictions of the described event. Each stimulus sentence is read twice.

Test B4--Sequence--Before. The child listens to a sentence with two events and then selects the one of two pictures that shows either the first or the last event. Picture order of the events follows their order of mention in the stimulus sentence. The answer choices cannot be ordered without the accompanying sentence. The two independent events are presented in clauses that are joined by the subordinate "before" in these sentence constructions: (1) Event 1 before Event 2; and (2) Before Event 2, Event 1. The two sentence constructions interchange the events in order of occurrence and

order of mention. Of the 20 sentences on the test, there are 10 of each type of sentence construction. Half of the 20 questions ask what happened first and half ask what happened last. These two question types are divided equally between the two sentence constructions. Item order for the sentence constructions and questions was randomly assigned. The test administrator reads each stimulus sentence twice and each question once.

Test B5--Sequence--After. This test has item types and proportions identical to those in Test B4, except the subordinate "after" is used in sentence constructions: (1) After Event 1, Event 2; and (2) Event 2 after Event 1.

Test C1--Detail: Negatives. The child answers 15 questions about specific details in five stories. The stories are written in simple, active voice sentences that have or do not have a negative; the questions do not paraphrase the story content. The parts of speech tested include objects of directional prepositions, pronouns, indirect objects, and objects of infinitives. Samples of the same part of speech being questioned are provided as the two or three incorrect answers for each item. The samples are either taken directly from the stories or represent a logical completion to the question stem.

Test C2--Paraphrase: Negatives. The child indicates whether 20 sets of sentences A and B have similar meanings by responding yes or no. The items are written in simple positive and simple negative sentences in the active voice and may include the coordinate "or" and a directional preposition.[2] A negative sentence contains such words as no, not, never, or negative verb

[2]Prepositions used in a directional sense usually modify action verbs, e.g., I am going toward the house.

forms. Sentence A is restated by rearranging the word order or by substituting synonyms to form Sentence B. Sentences A and B have similar meanings in 9 items and different meanings in 11 items.

Test C3--Central Thought--Topic: Without Organizer. The child identifies the topic of each of 15 passages without the aid of a passage organizer (topic sentence). There are two types of passages, dependent and independent, with approximately half of each per test form. (See footnote 1 for a definition of dependent and independent passages.) The topic, selected from three or four answer choices, is expressed in a single word or in a two- to four-word phrase. Incorrect choices are one- to three-word phrases about the details in the passage.

Test C4--Conclusions: One Relationship. The child uses the information in a short passage to determine whether a conclusion is deductively right or wrong. Each of the 20 passages includes a relationship sentence that is either class inclusive (e.g., "All the boats on the lake are sailboats.") or conditional (e.g., "If Patty goes to the store, she will ride her bike."), and a reference sentence that gives an instance or example of part of the relationship (e.g., "John's boat is on the lake; Patty is going to the store.") Half of the passages include a third sentence that gives information not useful in determining the answer.

Test C5--Sequence--Before/After. The child reads a sentence with two events and then determines which event occurred first or last. Answer choice order of the events follows their order of mention in the stimulus sentence. The two events do not have a natural, or dependent, sequence; their order can be determined only from the sentence. The two independent events are pre-

sented in clauses that are joined by the subordinates "before" or "after" in these sentence constructions: (1) Event 1 before Event 2; (2) After Event 1, Event 2; (3) Before Event 2, Event 1; and (4) Event 2 after Event 1. The four sentence constructions, which interchange the events in both order of occurrence and order of mention, are included in equal numbers on the test. Approximately half of the 20 questions ask what happened first and half ask what happened last. Item order for the sentence constructions and questions was randomly assigned.

Test D1--Context--Direct Clues. The child uses explicit context clues to select the best meaning for a nonsense word in a given sentence. One of four types of clues--synonym, definition, equivalent phrase, or summary--is given in a two- to five-sentence context for each item. In the 17-item test, summary clues are used in 8 items each and equivalent phrase, synonym, and definition clues are used in 3 items each. All parts of speech are tested. The correct answer is determined only through use of the explicit clues; no language experience clues are given. The two or three incorrect answers are the same part of speech as the nonsense word and are grammatically correct. One choice has a meaning opposite that of the correct answer, while the others present plausible answers.

Test D2--Detail: Passives. The child answers 20 questions about specific details in 5 stories. The stories are written in active and passive voice sentences which contain temporal or abstract prepositions and a dependent clause. The questions do not paraphrase story content. The parts of speech tested include objects of prepositions, objects of infinitives, relative pronouns, direct objects, indirect objects, and pronouns. Samples of the same part of speech being questioned are provided as the two or

three incorrect answers for each item. The samples are either taken directly from the stories or represent a logical completion to the question stem.

Test D3--Paraphrase: Passives. The child indicates whether 20 sets of sentences A and B have similar meanings by responding yes or no. The items are written in both active and passive voice sentences with no more than one subordinate clause. The sentences may include the coordinate "but," a causal subordinate, relative pronouns, and temporal or abstract prepositions.[3] Sentence A is restated by rearranging the word order and by substituting words (using synonyms) to form Sentence B. In half of the items the meanings of sentences A and B are similar and in the other half the meanings are different.

Test D4--Central Thought--Topic: Relevant Information. The child indicates whether three or four sentences from a passage are relevant or irrelevant to the passage topic. Relevant information is information that supports or describes the topic. Irrelevant information is information that does not support or describe the topic. There are two types of passages, dependent and independent, three of each per test form. (See Footnote1 for a definition of dependent and independent passages.) Each passage has an organizer (topic sentence) which is almost always the first sentence in the passage. The child responds to 20 sentences from the 6 passages. For each passage type, dependent or independent, 5 sentences are relevant to the topic and 5 are irrelevant.

---

[3]Prepositions used in the temporal sense describe time relationships (e.g., "Tom arrived about 10 p.m."). Prepositions used in the abstract sense describe relationships which are not concrete and are harder for children to determine (e.g., "Give me four quarters for a dollar").

Test D5--Relationships/Conclusions: Cause-Effect. After reading a short passage that has two cause-effect relationships, the child indicates whether the event in sentence A causes or does not cause the event in sentence B to happen. The events in sentences A and B are the cause or effect parts of the relationship sentence. The cause part of the relationship, introduced by the words "if," "when," "because," or "since," may appear before or after the effect part of the relationship. Each passage includes one to three sentences of information not useful for determining the answer. The 20-item test includes two sets of sentences A and B, one set for each relationship, for each of 10 passages.

Test D6--Sequence--Calendar Markers. The child determines the time relationship of two events in a two- or three-sentence passage. The two events do not have a natural, or dependent, sequence; their order can be determined only from the calendar marker clue given with each event. Five categories of clues test these relationships: (1) seasons to months (e.g., spring-April); (2) seasons to seasons (summer-fall); (3) specific days (Tuesday-Thursday); (4) whole days (yesterday-tomorrow); and (5) parts of days (morning-afternoon). The events in each passage are presented either in calendar order or out of order. Answer choice order for items testing category 1 is the calendar order of months; answer choice order for categories 2-5 follows the order of mention in the passage. The questions ask either when an event occurred or which event occurred first or last. Item order for the categories of clues and questions was randomly assigned.

Test E1--Word Parts--Prefixes. The student is provided with a list of 22 selected prefixes and their meanings to use as needed on either of the two parts of the test. Each part has five items. The specific prefixes tested

are representative of those on the list provided with the test. The list resulted from a frequency of use study which examined basal readers, content area materials, standardized tests, published word lists, and published research studies. On Part I the student adds a prefix to a base word to produce a word which expresses a given idea. The response is given full credit (2 points) if the prefix written is one that is traditionally combined with that base word to produce a word with that meaning. If the response produces a word which fits the given definition but is not a prefix traditionally combined with that base word, partial credit is given (1 point). On Part II the student reads a sentence in which one prefixed word is underlined and then writes a response about the meaning of the underlined word. Full credit (2 points) is given for either (1) a definition which synthesizes the meaning of the prefix with that of the base word or (2) a restatement of the stimulus sentence which demonstrates that the student has synthesized the meaning of the prefix and the base word. Partial credit is not given on Part II.

Test E2--Context--Direct Clues: Application. The student uses explicit context clues to derive a logical meaning for a nonsense word in one or more sentences in the first part of each test item and then applies that meaning to complete an unfinished sentence with the same nonsense word in the second part of the item. The nonsense word has the same meaning in both sets of sentences. One of four types of clues--synonym, definition, equivalent phrase, or summary--is given in each item. Definition and equivalent phrase clues are used in 5 items each, and synonym and summary clues are used in 2 items each in the 14-item test. All parts of speech are tested. The correct answer is determined only through use of the

explicit clues; no language experience clues are given. The three incorrect answers are the same part of speech as the nonsense word and are grammatically correct. One choice has the meaning opposite that of the correct answer while two choices present plausible answers.

Test E3--Detail: Clauses. The student answers 20 questions about specific details in four stories. The story sentences contain prepositional phrases and clauses introduced by temporal conjunctions; the questions do not paraphrase the story content. The student is required to identify relative pronouns in 25 percent of the items and previously tested parts of speech in a new context in 75 percent of the items. Samples of the same part of speech being questioned are provided as the two or three incorrect answers for each item. The samples are either taken directly from the stories or represent a logical completion to the question stem.

Test E4--Paraphrase: Complex Sentences. The student indicates whether 20 sets of sentences A and B have similar meanings by responding yes or no. The items are written in complex sentences that have either two subordinate clauses of which one is temporal, or one subordinate temporal clause and one embedded part[4] or compound verb. The sentences also include up to two prepositions which may be used positionally, directionally, temporally, or abstractly. Sentence A is restated by rearranging the word order or by substituting words (using synonyms) to form sentence B. Synthesis of information from a sentence's two clauses is emphasized. The correct answer is yes 9 times and no 11 times.

---

[4] word or clause between the subject and verb

Test E5--Central Thought: With Organizer. The student selects the sentence that best expresses the central thought of each of 15 passages. There are 2 types of passages, dependent (6 per test form) and independent (9 per test form.) (See Footnote 1 for a definition of dependent and independent passages.) Each passage has an organizer (topic sentence) which is the first sentence in approximately half of the passages and either the middle or the last sentence in the remainder. The central thought statement, selected from three or four answer choices, paraphrases the passage organizer. Incorrect choices are sentences that paraphrase details in the passage.

Test E6--Conclusions: Direct Relationships. The student indicates whether given conclusions are right or wrong based on the information in a passage. Each passage includes: (1) one to three relationships which may be descriptive, comparison-contrast, or cause-effect types of relationships; (2) reference sentences that give instances or examples of part of the relationship; and (3) information irrelevant to making the conclusions. Each relationship is stated within one sentence; each type of relationship is represented at least once in the five passages. The student responds to four conclusions for each of five passages.

Test E7--Sequence--Explicit Cues. The student selects from four choices the event that occurred before, after, or at the same time as another event based on the information presented in a passage. The explicit cue words used to connect the 8 to 10 events in a passage are "then," "next," "until," "later," "as soon as," "finally," "when," "during," "soon," "while," "meanwhile," "as," or "began." The student reads four passages and answers 15 questions. Wrong answer choices present other events in the passage.

Test F1--Word Parts--Suffixes. The student is provided with a list of 26 selected suffixes and their meanings to use as needed on either of the two parts of the test. Each part has five items. The specific suffixes tested are representative of those on the list provided with the test. The list derives from a frequency of use study which examined basal readers, content area materials, standardized tests, published word lists and published research studies. On Part I the student adds a suffix to a given base word to produce a different word which expresses a given idea. The response is given full credit (2 points) if the suffix written is one that is traditionally combined with that base word to produce a word with that meaning. If the response produces a word which fits the given definition but is not a suffix traditionally combined with that base word, partial credit is given (1 point). On Part II the student adds a suffix to a given base word to produce a different word which fits into the context of a sentence. The response is given full credit (2 points) if the suffix written is the one that is traditionally used to alter that base word to produce the correct grammatical form. Partial credit (1 point) is given to a response which creates an acceptable grammatical form but which is not the traditionally accepted suffix used with that base word.

Test F2--Context--Indirect Clues: Application. The student uses implicit context clues to derive a logical meaning of a nonsense word in one or more sentences in the first part of each test item, and then applies that meaning to complete an unfinished sentence with the same nonsense word in the second part of the item. The nonsense word has the same meaning in both sets of sentences. One of four types of clues--example, cause-effect, modifying phrase, or comparison-contrast--is given in each item. Example

and modifying phrase clues are used in 4 items, and cause-effect and comparison-contrast clues are used in 3 items in the 14-item test. All parts of speech are tested. The correct answer is determined only through use of the implicit clues; no language experience clues are given. The three incorrect answers are the same part of speech as the nonsense word and are grammatically correct. One choice has the meaning opposite that of the correct answer, while two choices present plausible answers.

Test F3--Paraphrase: Prepositions. The student indicates whether 20 sets of sentences A and B have similar meanings by responding yes or no. The items are written in complex sentences that have two dependent clauses and two or more prepositional phrases per sentence. Adversative coordinates are stressed. Sentence A is restated by rearranging the word order or by substituting words (using synonyms) to form sentence B. Synthesis of information from the main clause with that from both the centrally embedded part and the introductory or terminal clause is emphasized. The correct response is yes 9 times and no 11 times.

Test F4--Central Thought: Without Organizer. The student selects the sentence that best expresses the central thought of each of 14 passages without a passage organizer (topic sentence). There are 2 types of passages, dependent (6 per test form) and independent (8 per test form). (See footnote 1 for a definition of dependent and independent passages.) The central thought statement, selected from three or four answer choices, includes the passage topic and additional information that gives a more complete description of the topic. Incorrect choices are sentences that paraphrase details in the passage.

Test F5--Conclusions: Indirect Relationships. The student indicates whether given conclusions are right or wrong based on the information in a passage. Each passage includes: (1) two to four relationships which may be descriptive, comparison-contrast, or cause-effect types of relationships; (2) reference sentences that give instances or examples of parts of the relationship; and (3) information irrelevant to making the conclusions. Each relationship is stated indirectly, that is, the information is given in more than one sentence, so that the student must synthesize that information in order to understand the relationship. The student reads seven passages and responds to 20 conclusions.

Test F6--Sequence--Implicit Cues. The student indicates where additional events should be inserted on a given time line. Outcomes of previous events and preparations for following events are provided as the time line clues for the events to be inserted. There are four time lines and 12 total events to be placed. Seven time line points, A-G, are available as choices for all items.

Test G1--Word Parts--Combining Forms. The student is provided with a list of 21 selected combining forms and their meanings to use as needed on either of the two parts of the test. Each part has five items. The specific combining forms tested are representative of those on the list provided with the test. The list derives from a frequency of use study which examined basal readers, content area materials, standardized tests, published word lists,and published research studies. On Part I the student writes the meaning of the underlined combining form in a given word. One point is given for each correct response. On Part II the student writes a defini-

tion of a given word by synthesizing the meanings of the two combining forms in the word. A response is given full credit (2 points) if it shows a logical synthesis of the information contained in both combining forms, partial credit (1 point) if it only transfers the information for both combining forms from the list provided, or no credit if it either transfers the information for only one combining form from the list, or is illogical.

Test G2--Word Parts--Word Roots. The student is provided with a list of 13 selected word roots and their meanings to use as needed on either of the two parts of the test. Each part has five items. The specific word roots tested are representative of those on the list provided with the test. The list derives from a frequency of use study which examined basal readers, content area materials, standardized tests, published word lists, and published research studies. On Part I the student writes the word root for each of five given words. One point is given for each correct response. On Part II the student writes the word root which must be combined with a given word part to form the complete word as defined in the item. A response is given full credit (2 points) regardless of which spelling variation for the correct word root has been used. Partial credit is not given.

Test G3--Context--Uncommon Meanings. The student uses explicit and implicit context clues to select the obscure meaning of a familiar word used in a sentence. The clues sample from the explicit type and the implicit type. The correct answer for the 12 items is always the most unfamiliar meaning (determined by students' ratings of possible meanings). The three incorrect answers are other, more familiar definitions of the word.

Test G4--Paraphrase: Restatements. The student rewrites one long sentence into four or more sentences or rewrites three or four short sentences into one long sentence. The stimulus sentences contain sentence linkers (for example, "however") and are written in the sentence structures emphasized in the Levels C-F Paraphrase tests. To transform the long sentences, the restatement must be made in four or more short sentences in each of items 1-3. To transform the short sentences, the restatement must be made in one long sentence in each of items 4-10. Responses must be complete sentences and must include all of the ideas expressed in the stimulus sentences. Although the primary focus here is on separating and combining ideas among sentences, students are encouraged to use the reordering and substitution skills emphasized at Levels C-F. The maximum total test score is 10, one point for each item. The item responses are judged as either correct or incorrect; no partial credit is given.

Test G5--Central Thought: Generation. The student writes a central thought statement for each of seven passages. A central thought statement is a sentence that includes the passage topic and additional information that gives a more complete description of the topic. Three of the passages have a passage organizer (topic sentence) which is usually the first sentence in the passage, while four of the passages do not have an organizer. A response receives no credit if the topic is not stated, one point if only the topic is stated, and additional points if the topic plus information that provides a more complete description of the topic is stated. From two to four total points are possible for each central thought statement; the total possible score is 20.

Test G6--Conclusions: Supported/Unsupported. The student indicates whether given conclusions are supported or not supported based on the information in a passage. Each passage includes: (1) one to three relationships which may be descriptive or comparison-contrast types of relationships; (2) reference sentences that state which part of the relationship did or did not happen; and (3) information irrelevant to making the conclusions. Each relationship is stated either directly, in one sentence, or indirectly, in more than one sentence, so that the student must synthesize the information in those sentences in order to understand the relationship. Based on the information given, the conclusion is either supported (right) or unsupported (could be either right or wrong). The student reads six passages and responds to 21 conclusions.

Test G7--Conclusions: Modification. An individually administered performance test is provided for assessment of this skill. The student reads a passage and generates one or more conclusions if he or she decides that a conclusion can be made based on the information given. Then the student reads another passage and makes appropriate modifications if he or she decides that, based on the additional information, the initial conclusions could be augmented or changed. The four sets of passages are composed of descriptive, comparison-contrast, and cause-effect types of relationships, reference sentences, and information not needed to make or modify a conclusion. The relationships are either directly or indirectly stated (see Test G6 description). The initial conclusions are augmented in two passage sets, completely changed in one passage set, and are not modified in one passage set. The student receives one point for (1) recognizing that he or she has insufficient information to make or modify the conclu-

sions, (2) every correct conclusion, and (3) every correct modification of the first conclusion for a possible total score of 15.

Test G8--Sequence--Implied/Stated Events. The student reads a passage and then indicates where selected events from that passage should be placed on a given time line. Each passage is a story containing an average of 12 events. The questioned events are either implied or identified by explicit cue words in the passage. The explicit cue words include "before" and "after," the calendar markers assessed at Level D, and the cue words assessed at Level E. The student reads and places five events from each of three passages on the appropriate time line. Answer choices are the time line intervals which number from four to six. These intervals are such calendar markers as seasons, clock time, days of the week, and parts of the day (morning, afternoon, evening).

IV

RATIONALE FOR TEST CRITERIA[5]

Criteria for the WTRSD are specific to the kind of skill assessed and to the recommended use of test results. Specifications for content, usability, and psychometric characteristics evolved from the behavioral objectives that define the skills to be assessed, from the implementation guidelines for the Wisconsin Design, and from the following statement of purpose for the Design:

1. to identify and describe behaviorally the skills which appear to be essential for competence in reading;
2. to assess individual pupil's skill development status;
3. to manage instruction of children with different skill development needs;
4. to monitor each pupil's progress (Otto & Askov, 1974, p. 1).

Content Specifications

For tests of distinct basic skills, like those in the Design's outline of skills and objectives, content specifications include three points: (1) appropriate method of assessment of the behavior specified in the objective; (2) definition of assessment of the complete range of behavior specified in the objective; (3) guidelines for providing

---

[5] In this section the word master is used in the way it was used in the Word Attack and Study Skills technical manual (Stewart, 1976). Teachers implementing the Comprehension program will recognize that the master of a Comprehension test is instructed as an Applier of the Comprehension skill.

answer choices that relate to typical errors made by nonmasters of the skill. When the skill can be assessed only in reference to a context such as a sentence, a fourth point is included: The context must be simple enough so that using it does not require skills at or above the level of the skill to be assessed. An example of the content specifications may be useful in clarifying the above points and in relating them to the test descriptions (in section III).

First, a master of Comprehension Level E Skill 7 should be able to correctly determine the sequence from the cue words used in the stimulus paragraph, and not just be able to derive the sequence on the basis of logical ordering. For example, events without implicit order, such as riding a bike, going for a walk, and fishing would be embedded in a stimulus paragraph instead of events with implicit order, such as getting up, making breakfast, and going to school.

Second, complete range of behavior was specified where appropriate, such as with the Word Parts strand. One of the skills emphasized can be used for getting meaning from prefixes (Level E); however, appropriate prefixes[6] have been selected to correspond with students' general vocabulary development.

Third, a rationale for answer choices is needed. This rationale must not only identify as many as possible of the errors that a nonmaster of the skill would make, but also be sufficiently straightforward so that a master of the skill can scan the offered choices to select the one he or she knows to be correct. For the Level F Central Thought skill,

---

[6]anti-, bi, com-, de-, di-, dis-, en-, ex-, in-, inter-, mis-, non-, post-, pre-, pro-, re-, semi-, sub-, super-, tri-, un-.

the child determines which one of four statements best expresses the central thought of a paragraph with no organizer (topic sentence). Each of the three distracting statements pertains to only one or two details in the paragraph.

Fourth, a competent reader will apply basic skills to a wide variety of materials including texts, references, newspapers, and novels. Assessment of mastery of the skills, themselves, should be made independent of the context. This is not to say that using the skills in context is unimportant; however, until it is known that one has mastered the skills, assessment within more than the minimum amount and complexity of context would measure both the skill and its application. It would not identify a child's instructional needs related solely to the skill, itself. When skills themselves, rather than their application are emphasized, any context within which the skills are tested should be appropriate for application of lower-level skills.

Only when these four concerns about the content of the test can be agreed upon is it appropriate to use additional criteria in evaluating the instrument. Test results are to be used for instructional planning; teachers must be confident that test content reflects instructional content considered important for their students. Once relevance of the content has been established, additional test characteristics will be of interest and importance.

Usability Specifications

Certain logical requirements may be stated for tests that are to be used for instructional management and for monitoring the progress

of children. These requirements relate to the time and special training required to administer and score the tests. It was considered essential to keep the testing time for a single skill brief enough so that it does not consume a large proportion of the recommended instructional time. Two hours per week for a period of three weeks is the suggested allotment of time for developmental instruction in a single skill; the maximum possible time for administration of the test related to a single skill was arbitrarily set at 25 minutes with a target of 10 to 15 minutes. It should be noted, however, that all testing times are approximate based on the length of time required for nearly all children in an appropriate sample to finish the test. The tests are not timed; children are to work on a test until they have completed it. In these skill tests, there is to be no penalty for a methodical style of working.

Directions for both administration and scoring of the tests are developed and tried out with the goal of making them readily usable by teachers or aides with no special training beyond a study of the administrator's manual. A sample of 6 to 10 children for whom the skill is appropriate participate in the first test tryout. When the directions are clear and the testing task is simple, children ask few if any clarifying questions during a tryout test. The scoring of the objective tests is easy as raw scores are simply the number of test items correct; the only conversion of scores is to percent correct. Guidelines are provided for scoring of the generation tests.

Psychometric Characteristics

During development, content and usability specifications for the tests are defined and refined. Desirable quantitative characteristics

of the tests evolve in this process of definition and refinement. There is also an interplay of the logical and empirical characteristics of the tests: If a test meets the logical requirements well and does not meet the empirical requirements, it may indicate that additional clarification of the skill is needed.

Psychometric characteristics or numerically stated standards for the tests fall into two categories: those related to each instrument independently, and those related to parallel forms of the instruments.

For Each Instrument. If a skill has been clearly and discretely defined and a test developed which meets the logical content requirements for that skill, the test should first have relatively high (.75 - .80 or greater) internal-consistency reliability as an indicator of test homogeneity. Higher reliability is not required for two reasons: First, the .90 - .95 reliability expected of long, wide-range achievement tests is generally too much to expect for brief, focused instruments, such as the WTRSD. Second, although the decision made on the basis of these test results affects the individual student, it determines placement for approximately six hours of instruction rather than classification for an extended period of time.

Each of the items in a test should discriminate between masters and nonmasters. That is, children who would be identified as masters of the skill or those who have high test scores, should select the correct answer choice; nonmasters, or those who score low, should select one of the incorrect choices.

Frequently, one aspect of test revision after data have been collected is alteration or deletion of one or more of the answer choices. However, when a small total set of choices can be logically specified by content experts, as for the Word Part test described earlier, this traditional revision step is not necessary and is not desirable. Reducing the number of possible choices for each, say, Prefix test question could make the test slightly more complex in that each set of answer choices would appear to be different.

For Parallel Forms. When tests are used to manage instruction and to monitor the progress of individual students, at least some of the tests will be administered more than once to the same student. The second administration could occur immediately after skill instruction for a single skill, possibly as soon as three weeks after the first administration. When there is a high probability of reassessment, the use of parallel forms is desirable. Both forms of the test must meet the same content, usability, and psychometric requirements. The psychometric (classical test theory) requirements of the mean score, the same standard deviation, and the same internal-consistency reliability estimate can also logically be applied to the two forms.

For parallel forms of instruments that are part of an instructional management program such as the Design, additional requirements may be stated. The primary additional criterion is related to the decision made on the basis of the test results; in a high proportion of the cases, the mastery/nonmastery decision must be the same regardless of the form administered. Because users of the Design are free to

establish a cutoff score at 85 percent or 90 percent of the items correct as well as at the minimum recommended 80 percent, results from parallel forms of the WTRSD should yield the same cutoff decision in a high proportion of cases at any of the decision points.

Requirements concerning anticipated values of correlations between parallel forms and among the parallel forms of all tests at a level may also be derived. Certainly the correlation between parallel forms of the test for the same skill should be quite high (.50-.60). As with the criterion for internal-consistency reliability of a single instrument, the restriction of score range for short tests such as the WTRSD does not always permit the parallel forms correlation of .90 or greater that would be expected for long, wide-range achievement tests. Another requirement for parallel forms is that the correlations of any one test with all of the other tests at that level should be similar for its two forms. For example, the correlation of Form P of Test 2 at Level E with the Form Q tests at the same level should be similar to the correlations of Form Q Test 2 with the Form Q tests, and vice versa.

Summary

Content specifications and usability requirements may be stated as empirical criteria for the tests to meet; tests not meeting the stated conditions may indicate the need for reconsideration of the logically stated specifications. Use of the words may and reconsideration is intentional. Development of instruments of the highest quality possible for the intended use is dependent upon a close interrelationship between concerns related to the content and concerns related to the measurement itself. Both the content and the measurement characteristics must

be defined to build on what is relevant among the existing body of knowledge and to reject the familiar but irrelevant. Experts in both fields must be prepared to explain and defend their points of view and to be flexible in the application of criteria.

Specifications listed for the WTRSD were derived from information about the nature of the skills to be assessed and the anticipated use of the test results. The skills are important to reading competence and are independent; they may be assessed and taught directly, and then lat[illegible], reinforced through application in another context. It is appropriate to set the same criterion for Applier status for all learners, rather than to specify levels of competence or opportunities for experience that vary among learners. Results are used in planning instructional programs that permit each child to attain the skills he or she not have at an appropriate rate. For another type of objective or for another use, each criterion would have to be examined for appropriateness, and the total set of criteria reviewed for completeness.

The person who relates the statements in this section to the data presented in the following section will recognize that the WTRSD do not all meet or exceed all of the desirable criteria. In evaluating these instruments, the potential user of WTRSD must consider not only the criteria and the extent to which each test meets the various criteria, but also the characteristics of the students and staff who will be using the tests and the use of the test results.

V

SUMMARY STATISTICS FOR THE 1977 COMMERCIAL EDITION OF WTRSD COMPREHENSION

The data reported in this section are to be used as indicators of the quality of the WTRSD Comprehension for instructional planning and evaluation. Data are not intended to be used as norms for the tests. The WTRSD are to be used locally within an ongoing instructional program that has objectives like those stated for the Wisconsin Design skills. The district using the WTRSD for local accountability purposes will establish the local equivalent of norms and will relate initial status and subsequent improvement in skill status to the emphasis placed on those skills locally.

Complete levels, A-G, including parallel forms of the objective tests, were administered to children in grades 1-7 in the Monona Grove (Wisconsin) School District in the fall of 1978. School staff judged the appropriate instructional level for each child, and then a corresponding level of the Design was assigned. Assignments and testing procedures were carried out according to the recommendations in the Teacher's Planning Guide: Comprehension and the Test Administrator's Manual for that level.

Table 3

Test Statistics and Intercorrelations

WTRSD Comprehension Level A Form P

(N = 151)

		No. of Items	Mean Raw Score	% Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level A	1--Topic: Pictures	12	11.29	94.08	1.48	.76	.69
	2--Relationships	15	14.07	93.80	1.55	.71	.80
	3--First/Last Event	15	14.09	93.93	1.99	.84	.77

Test No.	1	2	3
1	-		
2	.48	-	
3	.48	.45	-

Table 4

Test Statistics and Intercorrelations

WTRSD Comprehension Level B Form P

(N = 124)

		No. of Items	Mean Raw Score	% Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level B	1--Detail	15	12.82	85.47	2.21	.71	1.15
	2--Topic: With Organizer	15	13.11	87.40	2.18	.73	1.10
	3--Outcomes	12	9.23	76.92	2.20	.63	1.28
	4--Before	20	18.35	91.75	2.64	.84	1.03
	5--After	20	16.31	81.55	3.68	.83	1.47

Test No.	1	2	3	4	5
1	-				
2	.63	-			
3	.42	.49	-		
4	.46	.37	.34	-	
5	.37	.27	.25	48	-

Table 5

Test Statistics and Intercorrelations

WTRSD Comprehension Level C Form P

(N = 126)

		No. of Items	Mean Raw Score	% Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level C	1--Detail: Negatives	15	13.52	90.13	2.19	.78	.99
	2--Paraphrase: Negatives	20	16.49	82.45	2.94	.72	1.53
	3--Topic: Without Organizer	15	12.90	86.00	2.31	.72	1.17
	4--Conclusions: One Relationship	20	16.29	81.45	3.32	.78	1.53
	5--Before/After	20	16.19	80.95	3.39	.78	1.56

Test No.	1	2	3	4	5
1	-				
2	.46	-			
3	.51	.56	-		
4	.55	.61	.65	-	
5	.73	.58	.49	.49	-

Table 6

Test Statistics and Intercorrelations

WTRSD Comprehension Level D Form P

(N = 148)

		No. of Items	Mean Raw Score	Mean % Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level D	1--Direct Clues	17	13.70	80.59	2.66	.67	1.48
	2--Detail: Passives	20	18.44	92.20	2.18	.74	1.08
	3--Paraphrase: Passives	20	17.02	85.10	2.58	.66	1.47
	4--Topic: Relevant Information	20	16.04	80.20	3.40	.79	1.53
	5--Cause-Effect	20	13.98	69.90	4.08	.81	1.73
	6--Calendar Markers	20	14.65	73.25	3.74	.78	1.72

Test No.	1	2	3	4	5	6
1	-					
2	.54	-				
3	.47	.45	-			
4	.32	.32	.44	-		
5	.36	.25	.39	.40	-	
6	.36	.41	.53	.48	.37	-

Table 7

Test Statistics and Intercorrelations

WTRSD Comprehension Level E Form P

(N = 127)

		No. of Items	Mean Raw Score	% Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level E	2--Direct Clues: Application	14	11.70	83.57	1.93	.60	1.18
	3--Detail: Clauses	20	19.03	95.15	1.83	.77	.85
	4--Paraphrase: Complex Sentences	20	18.17	90.85	2.07	.65	1.19
	5--Central Thought: With Organizer	15	12.76	85.07	2.54	.78	1.15
	6--Conclusions: Direct Relationships	20	15.76	78.80	3.36	.76	1.62
	7--Explicit Cues	15	11.62	77.47	3.37	.83	1.35

Test No.	2	3	4	5	6	7
2	-					
3	.58	-				
4	.49	.59	-			
5	.26	.35	.25	-		
6	.21	.17	.29	.08	-	
7	.33	.39	.29	.52	.45	-

Table 8

Test Statistics and Intercorrelations

WTRSD Comprehension Level F Form P

(N = 96)

		No. of Items	Mean Raw Score	Mean % Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level F	2--Indirect Clues: Application	14	9.61	68.64	2.50	.61	1.51
	3--Paraphrase: Prepositions	20	16.86	84.30	2.99	.76	1.42
	4--Central Thought: Without Organizer	14	10.97	78.36	2.66	.74	1.29
	5--Conclusions: Indirect Relationships	20	15.92	79.60	3.10	.71	1.63
	6--Implicit Cues	12	8.32	69.33	2.87	.77	1.31

Test No.	2	3	4	5	6
2	-				
3	.55	-			
4	.52	.60	-		
5	.57	.55	.63	-	
6	.49	.49	.56	.62	-

Table 9

Test Statistics and Intercorrelations

WTRSD Comprehension Level G Form P

(N = 180)

		No. of Items	Raw Score	% Corr.	Std. Dev.	Hoyt r	Std. error Measurement
Level G	3--Uncommon Meanings	12	11.23	93.58	1.12	.48	.78
	6--Conclusions: Supported/Unsupported	21	15.97	76.05	2.49	.53	1.67
	8--Implied/Stated Events	15	11.05	73.67	2.77	.69	1.48

Test No.	3	6	8
3	-		
6	.31	-	
8	.42	.43	-

Table 10

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level A

(N = 151)

Raw Score	Topic: Pictures	Relationships	First/Last Event
0	0	0	0
1	0	0	1
2	0	0	0
3	0	0	0
4	1	1	1
5	2	0	1
6	3	0	0
7	0	1	0
8	2	1	1
9	5	1	1
10	6	0	1
11	32	3	4
12	100	8	6
13		15	11
14		43	28
15		78	96

No. of Level A Skills Mastered	0	1	2	3
No. of Children	4	4	10	133

Table 11

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level B

(N = 124)

Raw Score	Detail	Topic: With Organizer	Outcomes	Before	After
0	1	1	1	1	2
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	1	0	0
4	0	0	2	0	0
5	0	1	4	0	0
6	0	0	7	0	0
7	2	0	9	0	0
8	2	2	12	0	3
9	1	3	22	1	1
10	9	4	25	1	3
11	13	10	27	0	1
12	18	13	14	1	4
13	20	21		0	6
14	28	36		3	12
15	30	33		6	12
16				5	9
17				9	10
18				13	20
19				34	19
20				50	22

No. of Level B Skills Mastered	0	1	2	3	4	5
No. of Children	1	7	15	25	36	40

Table 12

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level C

(N = 126)

Raw Score	Detail: Negatives	Paraphrase: Negatives	Topic: Without Organizer	Conclusions: One Relationship	Before/ After
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	1	0	0	0	0
5	1	0	1	0	1
6	1	0	1	0	0
7	2	0	3	0	0
8	1	1	3	2	3
9	0	3	7	2	3
10	6	3	4	4	3
11	6	0	11	9	5
12	7	7	8	2	4
13	14	9	17	10	8
14	30	7	37	7	8
15	57	9	34	11	9
16		16		9	10
17		10		11	14
18		23		12	17
19		21		25	24
20		17		22	17

No. of Level C Skills Mastered	0	1	2	3	4	5
No. of Children	10	9	13	12	29	53

Table 13

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level D

(N = 148)

Raw Scores	Direct Clues	Detail: Passives	Paraphrase: Passives	Topic: Relevant Information	Cause-Effect	Calendar Markers
0	0	0	0	1	1	1
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	1	0
4	0	0	0	0	0	0
5	2	0	0	0	1	1
6	1	0	1	0	2	2
7	1	1	1	1	4	3
8	3	0	0	2	10	4
9	4	0	1	2	6	5
10	7	2	1	4	9	3
11	9	0	0	8	7	9
12	15	0	3	3	9	9
13	20	1	9	9	7	14
14	19	5	5	11	9	10
15	21	2	12	14	18	17
16	28	13	14	13	21	19
17	18	12	27	21	13	17
18		15	25	18	12	9
19		33	29	21	8	16
20		64	20	20	10	9

No. of Level D Skills Mastered	0	1	2	3	4	5	6
No. of Children	6	11	17	22	35	28	29

Table 14

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level E

(N = 127)

Raw Score	Direct Clues: Application	Detail: Clauses	Paraphrase: Complex Sentences	Central Thought: With Organizer	Conclusions: Direct Relationships	Explicit Cues
0	0	0	0	2	2	4
1	0	0	0	0	0	0
2	1	0	0	0	0	0
3	0	0	0	0	0	2
4	0	0	0	0	0	1
5	0	0	0	1	0	1
6	1	0	0	0	0	2
7	1	0	0	1	0	4
8	4	0	0	4	0	2
9	8	0	0	3	2	4
10	16	1	1	6	3	15
11	16	0	1	11	4	12
12	25	2	2	10	4	14
13	39	3	0	16	12	24
14	16	1	6	55	10	23
15		1	4	18	11	19
16		0	8		20	0
17		4	9		19	0
18		9	24		17	0
19		35	34		9	0
20		71	38		14	0

No. of Level E Skills Mastered	0	1	2	3	4	5	6
No. of Children	3	4	10	13	26	23	48

Table 15

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level F

(N = 96)

Raw Scores	Indirect Clues: Application	Paraphrase: Prepositions	Central Thought: Without Organizer	Conclusions: Indirect Relationships	Implicit Cues
0	0	0	0	0	0
1	0	0	0	0	1
2	0	0	0	0	3
3	1	0	0	0	6
4	0	0	1	0	3
5	5	0	2	1	2
6	5	1	2	0	8
7	7	0	7	0	9
8	18	1	7	2	12
9	11	1	12	1	11
10	7	3	8	1	17
11	18	1	10	4	12
12	12	2	8	3	12
13	7	2	18	9	
14	5	6	21	6	
15		6		12	
16		11		6	
17		11		14	
18		15		18	
19		22		11	
20		14		8	

No. of Level F Skills Mastered	0	1	2	3	4	5
No. of Children	14	17	16	17	18	14

Table 16

Frequency Distributions of Raw Scores, Form P,

Number, and Percent Mastering From None to All Skills

(Mastery Criterion at 80%)

WTRSD Comprehension Level G

(N = 180)

Raw Score	Uncommon Meanings	Conclusions: Supported/Unsupported	Implied/ Stated Events
0	0	0	1
1	0	0	0
2	0	0	1
3	0	0	1
4	0	0	0
5	0	0	2
6	1	0	4
7	2	0	12
8	3	0	11
9	6	1	14
10	26	3	16
11	40	2	37
12	102	10	23
13		16	22
14		20	17
15		25	19
16		19	
17		27	
18		27	
19		21	
20		6	
21		3	

No. of Level G Skills Mastered	0	1	2	3
No. of Children	10	64	49	57

Table 17

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level A Form P

Proportion of Students	Topic: Pictures	Relationships	First/Last Event
100			
95-99	6	11	5
90-94	5	1	10
85-89	1	1	
80-84			
75-79		1	
70-74		1	
65-69			
60-64			
55-59			
50-54			
45-49			
40-44			
35-39			
30-34			
25-29			
20-24			
15-19			
10-14			
05-09			
00-04	*2*	*0*	*4*
None	*10*	*15*	*11*

Note: Distribution of omits shown in italics.

Table 18

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level B Form P

Proportion of Students	Detail	Topic: With Organizer	Outcomes	Before	After
100					
95-99	2	5		10	
90-94	4	4		5	5
85-89	4	1	3	1	4
80-84	2	2	3	3	3
75-79	1		2		2
70-74	1	2			3
65-69		1	2	1	2
60-64			1		1
55-59			1		
50-54					
45-49	1				
40-44					
35-39					
30-34					
25-29					
20-24					
15-19					
10-14					
05-09					
00-04	*15*	*15*	*12*	*20*	*20*
None					

Note: Distribution of omits shown in italics.

Table 19

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level C Form P

Proportion of Students	Detail: Negatives	Paraphrase: Negatives	Topic: Without Organizer	Conclusions: One Relationship	Before/After
100					
95-99	3	1	1	1	
90-94	6	3	5	6	2
85-89	2	5	3	1	4
80-84	3	5	4	2	7
75-79	1	2	1	3	1
70-74		2		4	4
65-69		1		3	2
60-64		1	1		
55-59					
50-54					
45-49					
40-44					
35-39					
30-34					
25-29					
20-24					
15-19					
10-14					
05-09					
00-04	*5*	*1*	*4*	*6*	*7*
None	*10*	*14*	*11*	*14*	*13*

Note: Distribution of omits shown in italics.

Table 20

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level D Form P

Proportion of Students	Direct Clues	Detail: Passives	Paraphrase: Passives	Topic: Relevant Information	Cause-Effect	Calendar Markers
100						
95-99		7	1	1		
90-94		6	3	4	1	1
85-89	4	6	5	4	2	3
80-84	7	1	7	2	4	3
75-79	3		3	2	1	
70-74	1		1	3	2	6
65-69	1			2	3	
60-64	1				2	4
55-59				1		1
50-54				1	2	
45-49					3	
40-44						1
35-39						
30-34						
25-29						
20-24						
15-19						
10-14						
05-09						
00-04	*7*	*7*	*10*	*20*	*20*	*20*
None	*10*	*13*	*10*			

Note: Distribution of omits shown in italics.

Table 21

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level E Form P

Proportion of Students	Direct Clues: Application	Detail: Clauses	Paraphrase: Complex Sentences	Central Thought: With Organizer	Conclusions: Direct Relationships	Explicit Cues
100						
95-99	3	15	3	1		
90-94	4	1	9	3	1	
85-89	1	4	6	7	5	3
80-84	2		2	1	6	3
75-79				1	3	5
70-74	2			1	2	1
65-69					1	3
60-64	1					
55-59					2	
50-54				1		
45-49	1					
40-44						
35-39						
30-34						
25-29						
20-24						
15-19						
10-14						
05-09						
00-04	*5*	*5*		*15*	*20*	*20*
None	*9*	*15*	*20*			

Note: Distribution of omits shown in italics.

Table 22

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level F Form P

Proportion of Students	Indirect Clues: Application	Paraphrase: Prepositions	Central Thought: Without Organizer	Conclusions: Indirect Relationships	Implicit Cues
100					
95-99		2			
90-94	2	4	3		
85-89	1	4	4	8	
80-84	1	7	1	3	3
75-79	1			4	4
70-74		2		3	
65-69	1		3		
60-64	2		2	1	1
55-59	3		1	1	1
50-54	3	1			3
45-49					
40-44					
35-39					
30-34					
25-29					
20-24					
15-19					
10-14					
05-09					
00-04	*7*			*1*	*3*
None	*7*	*20*	*14*	*19*	*9*

Note: Distribution of omits shown in italics.

Table 23

Frequency Distribution of Proportion Correct

and Proportion Omitting Each Item

WTRSD Comprehension Level G Form P

Proportion of Students	Uncommon Meanings	Conclusions: Supported/Unsupported	Implied/ Stated Events
100			
95-99	4		
90-94	6	5	1
85-89	2	6	3
80-84		1	2
75-79		2	1
70-74		1	2
65-69		2	1
60-64			4
55-59			1
50-54			
45-49		2	
40-44			
35-39		1	
30-34		1	
25-29			
20-24			
15-19			
10-14			
05-09			
00-04	*2*	*1*	*15*
None	*10*	*20*	

Note: Distribution of omits shown in italics.

Table 24

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level A Form P

r_{ptbis}	Topic: Pictures	Relationships	First/Last Event
≥1.00			
95-99			
90-94			
85-89			
80-84			
75-79		1	
70-74	1		1
65-69			3
60-64	3	2	3
55-59	3	2	3
50-54	1	2	3
45-49	2	4	1
40-44		1	
35-39		1	1
30-34		1	
25-29	1		
20-24			
15-19		1	
10-14			
05-09			
00-04		*4*	
-04-00			
-09-05	*1*	*1*	*6*
-14-10	*5*		*1*
-19-15	*1*	*3*	*1*
-24-20			*1*
-29-25	*3*	*4*	*5*
-34-30	*2*	*3*	*4*
-39-35	*2*	*3*	*3*
-44-40	*2*	*4*	*2*
-49-45	*5*	*2*	*2*
-54-50	*3*	*3*	*4*
-59-55	*1*	*1*	*5*
-64-60	*1*	*1*	
-69-65		*1*	
-74-70			
-79-75			
-84-80			
-89-85			
-94-90			
-99-95			
≤-1.00			

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 25

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level B Form P

r_{ptbis}	Detail	Topic: With Organizer	Outcomes	Before	After
≥1.00					
95-99					
90-94					
85-89					
80-84					
75-79					
70-74					
65-69				1	
60-64	1	1		2	3
55-59	1	2		3	4
50-54	3	3	3	9	4
45-49	4	4	3	2	2
40-44	3	3	4		4
35-39	3	1	1	3	
30-34	1	1	1		3
25-29					
20-24					
15-19					
10-14					
05-09			*1*		
00-04	*10*	*7*	*4*		
-04-00	*2*	*5*	*1*		
-09-05	*4*	*4*	*5*	*2*	*1*
-14-10	*5*	*5*	*2*	*3*	*1*
-19-15	*9*	*5*	*6*	*1*	*3*
-24-20	*6*	*8*	*3*	*1*	*2*
-29-25	*2*	*6*	*6*	*1*	*1*
-34-30	*4*	*2*	*6*	*5*	*4*
-39-35	*5*	*3*	*13*	*3*	*4*
-44-40	*2*	*1*		*1*	*2*
-49-45			*1*	*3*	*3*
-54-50	*10*	*15*		*1*	*2*
-59-55					*17*
-64-60				*19*	
-69-65					
-74-70					
-79-75					
-84-80					
-89-85					
-94-90					
-99-95					
≤-1.00					

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 26

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level C Form P

r_{ptbis}	Detail: Negatives	Paraphrase: Negatives	Topic: Without Organizer	Conclusions: One Relationship	Before/After
≥1.00					
95-99					
90-94					
85-89					
80-84					
75-79	1				
70-74					
65-69	2			1	
60-64				2	
55-59	3	1	3	1	
50-54	3		2	1	6
45-49	2	6	1	3	6
40-44		3	7	3	2
35-39		1	2	2	3
30-34		4		5	3
25-29	1	3		1	
20-24		1			
15-19		1		1	
10-14					
05-09	1	*1*		*2*	
00-04	*17*		*6*	*2*	
-04-00	*3*				
-09-05	*1*		*6*	*1*	*5*
-14-10	*4*		*1*	*1*	
-19-15	*1*	*1*	*5*	*1*	*2*
-24-20	*1*	*2*	*9*		
-29-25	*4*	*2*	*7*	*1*	*1*
-34-30	*6*	*4*	*7*	*5*	*2*
-39-35	*1*	*1*	*4*	*2*	*3*
-44-40	*2*	*3*	*3*	*3*	*2*
-49-45	*7*	*6*	*1*	*2*	*6*
-54-50	*1*			*2*	*6*
-59-55	*1*	*1*		*1*	
-64-60	*2*			*2*	
-69-65	*1*			*1*	
-74-70					
-79-75					
-84-80					
-89-85					
-94-90					
-99-95					
≤-1.00					

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 27

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level D Form P

r_{ptbis}	Direct Clues	Detail: Passives	Paraphrase: Passives	Topic: Relevant Information	Cause-Effect	Calendar Markers
≥1.00						
95-99						
90-94						
85-89						
80-84						
75-79						
70-74						
65-69						
60-64		2	1	2	3	1
55-59	1	2	1	3	3	2
50-54	1	4	1	3	3	3
45-49	3	1	2		4	4
40-44	4	4	3	6	1	3
35-39	3	2	4	4	4	3
30-34	4	1	2		2	3
25-29	1	1	3	1		
20-24			3	1		
15-19		1				1
10-14		1				
05-09		1				
00-04	*7*	*23*	*1*			*16*
-04-00	*7*	*5*	*1*			
-09-05	*6*	*3*	*2*			
-14-10	*2*	*2*		*1*		*2*
-19-15	*6*	*5*	*1*	*1*	*2*	*1*
-24-20	*8*	*2*	*10*	*6*	*2*	*8*
-29-25	*8*	*4*	*2*	*8*	*11*	*7*
-34-30	*9*	*9*	*1*	*11*	*5*	*11*
-39-35	*4*	*3*	*5*	*6*	*8*	*8*
-44-40	*1*	*3*	*2*	*1*	*4*	*3*
-49-45		*4*	*3*	*3*	*4*	*2*
-54-50		*3*	*1*	*3*	*3*	*1*
-59-55			*1*		*1*	*1*
-64-60		*1*				
-69-65						
-74-70						
-79-75						
-84-80						
-89-85						
-94-90						
-99-95						
≤-1.00						

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 28

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level E Form P

r_{ptbis}	Direct Clues: Application	Detail: Clauses	Paraphrase: Complex Sentences	Central Thought: With Organizer	Conclusions: Direct Relationships	Explicit Cues
≥1.00						
95-99						
90-94						
85-89						
80-84						
75-79						
70-74		1				
65-69		1		2		1
60-64		1		1		2
55-59		2		3	1	5
50-54	2	1		2	1	3
45-49	2	6	6	3	6	4
40-44	5	3	1	3	6	
35-39	3	2	5		6	
30-34			5	1		
25-29	1		1			
20-24	1					
15-19		1	2			
10-14						
05-09		2				
00-04	*3*	*22*		*7*		*2*
-04-00	*5*	*5*		*4*		*3*
-09-05	*4*	*2*		*2*		*5*
-14-10		*2*		*10*		*7*
-19-15	*8*	*1*	*2*	*5*	*4*	*12*
-24-20	*11*	*7*		*10*	*7*	*9*
-29-25	*5*	*4*	*1*	*3*	*6*	*2*
-34-30	*3*	*5*	*5*	*3*	*2*	*4*
-39-35	*6*	*4*	*5*		*1*	*1*
-44-40	*1*	*7*	*1*	*1*		
-49-45		*3*	*6*			
-54-50	*1*	*1*		*3*		
-59-55				*3*	*20*	
-64-60				*9*		*15*
-69-65		*2*				
-74-70						
-79-75						
-84-80						
-89-85						
-94-90						
-99-95						
≤-1.00						

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 29

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level F Form P

r_{ptbis}	Indirect Clues: Application	Paraphrase: Prepositions	Central Thought: Without Organizer	Conclusions: Indirect Relationships	Implicit Cues
≥1.00					
95-99					
90-94					
85-89					
80-84					
75-79			1		
70-74					
65-69			1		
60-64			2	1	3
55-59	2	2	1	2	3
50-54	3	5	1	3	4
45-49	1	4	1	2	
40-44	3	2	2	3	1
35-39		4	3	3	1
30-34	2	1	1	2	
25-29		2			
20-24	2			3	
15-19			1		
10-14	1			1	*1*
05-09	*2*		*1*		
00-04			*3*		*21*
-04-00	*2*				*4*
-09-05	*4*	*1*	*2*	*1*	*3*
-14-10	*7*		*5*	*1*	*5*
-19-15	*17*		*3*		*9*
-24-20	*4*		*8*	*2*	*7*
-29-25	*3*	*2*	*6*		*13*
-34-30	*6*	*1*	*5*	*2*	*5*
-39-35	*5*	*4*	*4*	*4*	*6*
-44-40		*2*	*3*	*3*	
-49-45		*4*	*1*	*2*	*1*
-54-50		*5*		*3*	
-59-55		*2*		*2*	
-64-60				*1*	
-69-65			*1*		
-74-70					
-79-75					
-84-80					
-89-85					
-94-90					
-99-95					
≤-1.00					

Note: Distribution of correlations for incorrect alternatives s own in italics.

Table 30

Frequency Distribution of Item-Test Correlations (point biserial r)

WTRSD Comprehension Level G Form P

r_{ptbis}	Uncommon Meanings	Conclusions: Supported/Unsupported	Implied/ Stated Events
≥1.00			
95-99			
90-94			
85-89			
80-84			
75-79			
70-74			
65-69			
60-64			2
55-59	1		1
50-54	2		2
45-49	2	3	2
40-44		1	2
35-39	2	4	2
30-34	1	2	4
25-29	3	7	
20-24	1	3	
15-19			
10-14			
05-09		1	
00-04	*4*		*20*
-04-00	*5*		*7*
-09-05	*8*	*1*	*8*
-14-10		*1*	*20*
-19-15	*1*		*16*
-24-20	*3*	*3*	*14*
-29-25	*8*	*7*	*8*
-34-30	*3*	*2*	*9*
-39-35	*2*	*4*	*3*
-44-40	*2*	*2*	
-49-45	*2*	*2*	
-54-50			
-59-55			
-64-60			
-69-65			
-74-70			
-79-75			
-84-80			
-89-85			
-94-90			
-99-95			
≤-1.00			

Note: Distribution of correlations for incorrect alternatives shown in italics.

Table 31

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level A

(N = 55)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
1--Topic: Pictures	P	54	98.18	.96		
	Q	53	96.36	.98	94.55	.03
2--Relationships	P	55	100.00	1.00		
	Q	55	100.00	1.00	100.00	--[a]
3--First/Last Event	P	54	98.18	1.00		
	Q	55	100.00	.98	98.18	--[a]
Criterion at 85%						
1--Topic: Pictures	P	53	96.36	.94		
	Q	52	94.55	.96	90.91	.05
2--Relationships	P	55	100.00	.96		
	Q	53	96.36	1.00	96.36	--[a]
3--First/Last Event	P	53	96.36	1.00		
	Q	55	100.00	.96	96.36	--[a]
Criterion at 90%						
1--Topic: Pictures	P	53	96.36	.94		
	Q	52	94.55	.96	90.91	.05
2--Relationships	P	54	98.18	.94		
	Q	52	94.55	.98	92.73	.03
3--First/Last Event	P	51	92.73	.94		
	Q	51	92.73	.94	87.27	.19

[a] Phi cannot be computed.

Table 32

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level B

(N = 37)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
1--Detail	P	30	81.08	.93		
	Q	33	89.19	.85	75.68	.28
2--Topic: With Organizer	P	35	94.59	.97		
	Q	35	94.59	.97	91.89	.47
3--Outcomes	P	23	62.16	.87		
	Q	27	72.97	.74	54.05	.40
4--Before	P	33	89.19	1.00		
	Q	36	97.30	.92	89.19	.48
5--After	P	26	70.27	.92		
	Q	31	83.78	.77	64.86	.36
Criterion at 85%						
1--Detail	P	30	81.08	.83		
	Q	30	81.08	.83	67.57	.12
2--Topic: With Organizer	P	33	89.19	.91		
	Q	33	89.19	.91	81.08	.16
3--Outcomes	P	15	40.54	.53		
	Q	17	45.95	.47	21.62	.12
4--Before	P	31	83.78	1.00		
	Q	32	86.49	.97	83.78	.90
5--After	P	24	64.86	.92		
	Q	29	78.38	.76	59.46	.44
Criterion at 90%						
1--Detail	P	23	62.16	.78		
	Q	26	70.27	.69	48.65	.22
2--Topic: With Organizer	P	27	72.97	.81		
	Q	29	78.38	.76	59.46	.12
3--Outcomes	P	15	40.54	.53		
	Q	17	45.95	.47	21.62	.12
4--Before	P	30	81.08	.97		
	Q	31	83.78	.94	78.38	.72
5--After	P	20	54.05	.80		
	Q	24	64.86	.67	43.24	.34

Table 33

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level C

(N = 40)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
1--Detail: Negatives	P	40	100.00	1.00		
	Q	40	100.00	1.00	100.00	--[a]
2--Paraphrase: Negatives	P	37	92.50	.86		
	Q	33	82.50	.97	80.00	.37
3--Topic: Without Organizer	P	39	97.50	.97		
	Q	39	97.50	.97	95.00	.03
4--Conclusions: One Relationship	P	33	82.50	.64		
	Q	25	62.50	.84	52.50	.05
5--Before/After	P	31	77.50	.94		
	Q	34	85.00	.85	72.50	.44
Criterion at 85%						
1--Detail: Negatives	P	37	92.50	1.00		
	Q	40	100.00	.92	92.50	--[a]
2--Paraphrase: Negatives	P	32	80.00	.91		
	Q	32	80.00	.91	72.50	.53
3--Topic: Without Organizer	P	34	85.00	.94		
	Q	35	87.50	.91	80.00	.48
4--Conclusions: One Relationship	P	31	77.50	.61		
	Q	24	60.00	.79	47.50	.05
5--Before/After	P	26	65.00	.81		
	Q	28	70.00	.75	52.50	.32
Criterion at 90%						
1--Detail: Negatives	P	33	82.50	.82		
	Q	33	82.50	.82	67.50	.04
2--Paraphrase: Negatives	P	26	65.00	.77		
	Q	27	67.50	.74	50.00	.27
3--Topic: Without Organizer	P	29	72.50	.93		
	Q	31	77.50	.87	67.50	.61
4--Conclusions: One Relationship	P	27	67.50	.44		
	Q	17	42.50	.71	30.00	.06
5--Before/After	P	23	57.50	.65		
	Q	18	45.00	.83	37.50	.47

[a] Phi cannot be computed.

Table 34

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level D

(N = 36)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
1--Direct Clues	P	24	66.67	.88		
	Q	29	80.56	.72	58.33	.25
2--Detail: Passives	P	35	97.22	.97		
	Q	35	97.22	.97	94.44	.03
3--Paraphrase: Passives	P	31	86.11	.81		
	Q	27	75.00	.93	69.44	.32
4--Topic: Relevant Information	P	26	72.22	.81		
	Q	24	66.67	.88	58.33	.48
5--Cause-Effect	P	18	50.00	.72		
	Q	21	58.33	.62	36.11	.28
6--Calendar Markers	P	23	63.89	.87		
	Q	25	69.44	.80	55.56	.51
Criterion at 85%						
1--Direct Clues	P	17	47.22	.59		
	Q	22	61.11	.45	27.78	.04
2--Detail: Passives	P	35	97.22	.94		
	Q	34	94.44	.97	91.67	.04
3--Paraphrase: Passives	P	29	80.56	.76		
	Q	24	66.67	.92	61.11	.40
4--Topic: Relevant Information	P	23	63.89	.70		
	Q	21	58.33	.76	44.44	.30
5--Cause-Effect	P	15	41.67	.53		
	Q	15	41.67	.53	22.22	.20
6--Calendar Markers	P	15	41.67	.60		
	Q	17	47.22	.53	25.00	.22
Criterion at 90%						
1--Direct Clues	P	12	33.33	.42		
	Q	14	38.89	.36	13.89	.04
2--Detail: Passives	P	34	94.44	.91		
	Q	33	91.67	.94	86.11	.07
3--Paraphrase: Passives	P	24	66.67	.79		
	Q	21	58.33	.90	52.78	.60
4--Topic: Relevant Information	P	20	55.56	.55		
	Q	15	41.67	.73	30.56	.30
5--Cause-Effect	P	12	33.33	.33		
	Q	10	27.78	.40	11.11	.09
6--Calendar Markers	P	10	27.78	.50		
	Q	9	25.00	.56	13.89	.36

Table 35

Parallel Forms Mastery-Decision Characteristics
WTBSD Comprehension Level E
(N = 50)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
2--Direct Clues: Application	P	32	64.00	.75		
	Q	29	58.00	.83	48.00	.46
3--Detail: Clauses	P	46	92.00	.98		
	Q	48	96.00	.94	90.00	.32
4--Paraphrase: Complex Sentences	P	46	92.00	.96		
	Q	46	92.00	.96	88.00	.46
5--Central Thought: With Organizer	P	40	80.00	.92		
	Q	42	84.00	.88	74.00	.46
6--Conclusions: Direct Relationships	P	34	68.00	.85		
	Q	39	78.00	.74	58.00	.26
7--Explicit Cues	P	31	62.00	.58		
	Q	24	48.00	.75	36.00	.26
Criterion at 85%						
2--Direct Clues: Application	P	32	64.00	.75		
	Q	29	58.00	.83	48.00	.46
3--Detail: Clauses	P	46	92.00	.98		
	Q	47	94.00	.96	90.00	.55
4--Paraphrase: Complex Sentences	P	44	88.00	.86		
	Q	40	80.00	.95	76.00	.43
5--Central Thought With Organizer	P	36	72.00	.89		
	Q	39	78.00	.82	64.00	.42
6--Conclusions: Direct Relationships	P	26	52.00	.65		
	Q	33	66.00	.52	34.00	.01
7--Explicit Cues	P	26	52.00	.58		
	Q	19	38.00	.79	30.00	.42
Criterion at 90%						
2--Direct Clues: Application	P	20	40.00	.85		
	Q	27	54.00	.63	34.00	.51
3--Detail: Clauses	P	42	84.00	.93		
	Q	43	86.00	.91	78.00	.45
4--Paraphrase: Complex Sentences	P	39	78.00	.64		
	Q	28	56.00	.89	50.00	.31
5--Central Thought: With Organizer	P	26	52.00	.77		
	Q	30	60.00	.67	40.00	.36
6--Conclusions: Direct Relationships	P	19	38.00	.63		
	Q	26	52.00	.46	24.00	.17
7--Explicit Cues	P	20	40.00	.45		
	Q	15	30.00	.60	18.00	.27

Table 36

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level F

(N = 46)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
2--Indirect Clues: Application	P	11	23.91	.45		
	Q	11	23.91	.45	10.87	.28
3--Paraphrase: Prepositions	P	38	82.61	.76		
	Q	34	73.91	.85	63.04	.12
4--Central Thought: Without Organizer	P	17	36.96	.71		
	Q	28	60.87	.43	26.09	.15
5--Conclusions: Indirect Relationships	P	19	41.30	.63		
	Q	16	34.78	.75	26.09	.50
6--Implicit Cues	P	15	32.61	.13		
	Q	4	8.70	.50	4.35	.11
Criterion at 85%						
2--Indirect Clues: Application	P	11	23.91	.45		
	Q	11	23.91	.45	10.87	.28
3--Paraphrase: Prepositions	P	32	69.57	.66		
	Q	29	63.04	.72	45.65	.08
4--Central Thought: Without Organizer	P	17	36.96	.71		
	Q	28	60.87	.43	26.09	.15
5--Conclusions: Indirect Relationships	P	14	30.43	.57		
	Q	10	21.74	.80	17.39	.57
6--Implicit Cues	P	9	19.57	.11		
	Q	2	4.35	.50	2.17	.16
Criterion at 90%						
2--Indirect Clues: Application	P	4	8.70	0.00		
	Q	4	8.70	0.00	0.00	.10
3--Paraphrase: Prepositions	P	27	58.70	.48		
	Q	21	45.65	.62	28.26	.06
4--Central Thought: Without Organizer	P	13	28.26	.77		
	Q	22	47.83	.45	21.74	.37
5--Conclusions: Indirect Relationships	P	10	21.74	.40		
	Q	6	13.04	.67	8.70	.42
6--Implicit Cues	P	9	19.57	.11		
	Q	2	4.35	.50	2.17	.16

Table 37

Parallel Forms Mastery-Decision Characteristics

WTRSD Comprehension Level G

(N = 50)

	Form	Masters No.	Masters %	Probability Master will also master other form	Proportion of cases with same mastery decisions	ϕ
Criterion at 80%						
3--Uncommon Meanings	P	44	88.00	.95		
	Q	44	88.00	.95	84.00	.62
6--Conclusions: Supported/ Unsupported	P	23	46.00	.65		
	Q	22	44.00	.68	30.00	.39
8--Implied/ Stated Events	P	16	32.00	.56		
	Q	18	36.00	.50	18.00	.29
Criterion at 85%						
3--Uncommon Meanings	P	37	74.00	.89		
	Q	38	76.00	.87	66.00	.52
6--Conclusions: Supported/ Unsupported	P	15	30.00	.53		
	Q	15	30.00	.53	16.00	.33
8--Implied/ Stated Events	P	13	26.00	.46		
	Q	14	28.00	.43	12.00	.24
Criterion at 90%						
3--Uncommon Meanings	P	37	74.00	.89		
	Q	38	76.00	.87	66.00	.52
6--Conclusions: Supported/ Unsupported	P	7	14.00	.71		
	Q	12	24.00	.42	10.00	.45
8--Implied/ Stated Events	P	7	14.00	.29		
	Q	8	16.00	.25	4.00	.14

Table 38

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level A, Forms P & Q

(N = 55)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	1--Topic: Pictures	11.76	98.00	.5762
P	2--Relationships	14.60	97.33	.5305
P	3--First/Last Event	14.65	97.67	.7986
Q	1--Topic: Pictures	11.69	97.42	.7667
Q	2--Relationships	14.51	96.73	.7168
Q	3--First/Last Event	14.78	98.53	.5673

Table 39

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level B, Forms P & Q

(N = 37)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	1--Detail	13.46	89.73	1.5201
P	2--Topic: With Organizer	13.97	93.13	1.4237
P	3--Outcomes	9.65	80.42	2.0712
P	4--Before	18.49	92.45	1.6770
P	5--After	17.00	85.00	3.2318
Q	1--Detail	13.81	92.07	1.4306
Q	2--Topic: With Organizer	14.03	93.53	1.1177
Q	3--Outcomes	9.78	81.50	1.9598
Q	4--Before	19.00	95.00	1.4720
Q	5--After	17.73	88.65	2.7249

Table 40

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level C, Forms P & Q

(N = 40)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	1--Detail: Negatives	14.32	95.47	.9443
P	2--Paraphrase: Negatives	17.88	89.40	2.1023
P	3--Topic: Without Organizer	13.88	92.53	1.2848
P	4--Conclusions: One Relationship	17.55	87.75	2.3637
P	5--Before/After	16.97	84.85	2.9655
Q	1--Detail: Negatives	14.40	96.00	.7779
Q	2--Paraphrase: Negatives	17.47	87.35	2.6697
Q	3--Topic: Without Organizer	14.17	94.47	1.2171
Q	4--Conclusions: One Relationship	16.42	82.10	2.9862
Q	5--Before/After	17.02	85.10	2.3037

Table 41

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level D, Forms P & Q

(N = 36)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	1--Direct Clues	14.36	84.47	1.7263
P	2--Detail: Passives	19.44	97.20	1.2971
P	3--Paraphrase: Passives	18.03	90.15	1.6645
P	4--Topic: Relevant Information	16.64	83.20	3.3819
P	5--Cause-Effect	15.19	75.95	3.9120
P	6--Calendar Markers	15.69	78.45	2.7131
Q	1--Direct Clues	14.58	85.76	2.5228
Q	2--Detail: Passives	19.03	95.15	2.5800
Q	3--Paraphrase: Passives	17.25	86.25	2.8322
Q	4--Topic: Relevant Information	15.92	79.60	3.4918
Q	5--Cause-Effect	14.92	74.60	3.6440
Q	6--Calendar Markers	16.06	80.30	2.1639

Table 42

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level E, Forms P & Q

(N = 50)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	2--Direct Clues: Application	11.88	84.86	1.6980
P	3--Detail: Clauses	18.74	93.70	2.4312
P	4--Paraphrase: Complex Sentences	18.34	91.70	1.8251
P	5--Central Thought: With Organizer	12.72	84.80	2.4993
P	6--Conclusions: Direct Relationships	16.12	80.60	2.8971
P	7--Explicit Cues	11.74	78.27	3.0089
Q	2--Direct Clues: Application	11.86	84.71	2.4160
Q	3--Detail: Clauses	18.86	94.30	1.3704
Q	4--Paraphrase: Complex Sentences	17.62	88.10	2.1082
Q	5--Central Thought: With Organizer	13.04	86.93	2.8064
Q	6--Conclusions: Direct Relationships	16.76	83.80	3.2170
Q	7--Explicit Cues	11.16	74.40	3.0796

Table 43

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level F, Forms P & Q

(N = 46)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	2--Indirect Clues: Application	9.52	68.00	2.3261
P	3--Paraphrase: Prepositions	17.41	87.05	2.0284
P	4--Central Thought: Without Organizer	10.63	75.93	2.4527
P	5--Conclusions: Indirect Relationships	15.17	75.85	2.5500
P	6--Implicit Cues	8.11	67.58	2.7262
Q	2--Indirect Clues: Application	9.41	67.21	2.4996
Q	3--Paraphrase: Prepositions	16.70	83.50	2.7069
Q	4--Central Thought: Without Organizer	11.30	80.71	2.8116
Q	5--Conclusions: Indirect Relationships	13.83	69.15	3.1220
Q	6--Implicit Cues	6.43	53.58	2.6470

Table 44

Sample Characteristics for Intersubtest Analysis

WTRSD Comprehension Level G, Forms P & Q

(N = 50)

Form	Test	Mean Raw Score	Mean % Corr.	Standard Deviation
P	3--Uncommon Meanings	11.04	92.00	1.2115
P	6--Conclusions: Supported/ Unsupported	15.[illegible]	74.29	3.0772
P	8--Implied/ Stated Events	10.14	67.60	3.1103
Q	3--Uncommon Meanings	11.00	91.67	1.3702
Q	6--Conclusions: Supported/ Unsupported	15.88	75.62	3.0947
Q	8--Implied/ Stated Events	9.24	61.60	3.8943

Table 45

Intersubtest Conditional Probabilities for Mastering Tests,

Criterion at 80%

WTRSD Comprehension Level A, Forms P & Q

(N = 55)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability											
		N	Topic: Pictures P	N	Relation-ships P	N	First/ Last Event P	N	Topic: Pictures Q	N	Relation-ships Q	N	First/ Last Event Q
Topic: Pictures--P	A	54	1.000	54	1.000	54	.981	54	.963	54	1.000	54	1.000
	NA	1	0.000	1	1.000	1	1.000	1	1.000	1	1.000	1	1.000
Relation-ships--P	A	55	.982	55	1.000	55	.982	55	0.964	55	1.000	55	1.000
	NA	0	0.000	0	0.000	0	0.000	0	.000	0	0.000	0	0.000
First/Last Event--P	A	54	.981	54	1.000	54	1.000	54	.963	54	1.000	54	1.000
	NA	1	1.000	1	1.000	1	0.000	1	1.000	1	1.000	1	1.000
Topic: Pictures--Q	A	53	.981	53	1.000	53	.981	53	1.000	53	1.000	53	1.000
	NA	2	1.000	2	1.000	2	1.000	2	0.000	2	1.000	2	1.000
Relation-ships--Q	A	55	.982	55	1.000	55	.982	55	.964	55	1.000	55	1.000
	NA	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
First/Last Event--Q	A	55	.982	55	1.000	55	.982	55	.964	55	1.000	55	1.000
	NA	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000

Table 46

Intersubtest Conditional Probabilities for Mastering Tests,

Criterion at 80%

WTRSD Comprehension Level B, Forms P & Q

(N = 37)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability																	
		N	Detail P	N	Topic: With Organizer P	N	Outcomes P	N	Before P	N	After P	N	Detail Q	N	Topic: With Organizer Q	N	Outcomes Q	N	Before Q	N	After Q
Detail--P	A	30	1.000	30	.933	30	.700	30	.933	30	.767	30	.933	30	.933	30	.800	30	1.000	30	.867
	NA	7	0.000	7	1.000	7	.286	7	.714	7	.429	7	.714	7	1.000	7	.429	7	.857	7	.714
Topic: With Organizer--P	A	35	.800	35	1.000	35	.657	35	.886	35	.743	35	.886	35	.971	35	.771	35	.971	35	.857
	NA	2	1.000	2	0.000	2	0.000	2	1.000	2	0.000	2	1.000	2	.500	2	0.000	2	1.000	2	.500
Outcomes--P	A	23	.913	23	1.000	23	1.000	23	.957	23	.783	23	.826	23	.957	23	.870	23	1.000	23	.870
	NA	14	.643	14	.857	14	0.000	14	.786	14	.571	14	1.000	14	.929	14	.500	14	.929	14	.786
Before--P	A	33	.848	33	.939	33	.667	33	1.000	33	.758	33	.879	33	.939	33	.758	33	1.000	33	.848
	NA	4	.500	4	1.000	4	.250	4	0.000	4	.250	4	1.000	4	1.000	4	.500	4	.750	4	.750
After--P	A	26	.885	26	1.000	26	.692	26	.962	26	1.000	26	.846	26	.962	26	.846	26	1.000	26	.923
	NA	11	.636	11	.818	11	.455	11	.727	11	0.000	11	1.000	11	.909	11	.455	11	.909	11	.636
Detail--Q	A	33	.848	33	.939	33	.576	33	.879	33	.667	33	1.000	33	.939	33	.727	33	.970	33	.818
	NA	4	.500	4	1.000	4	1.000	4	1.000	4	1.000	4	0.000	4	1.000	4	.750	4	1.000	4	1.000
Topic: With Organizer--Q	A	35	.800	35	.971	35	.629	35	.886	35	.714	35	.886	35	1.000	35	.743	35	.971	35	.857
	NA	2	1.000	2	.500	2	.500	2	1.000	2	.500	2	1.000	2	0.000	2	.500	2	1.000	2	.500
Outcomes--Q	A	27	.889	27	1.000	27	.741	27	.926	27	.815	27	.889	27	.963	27	1.000	27	.963	27	.889
	NA	10	.600	10	.800	10	.300	10	.800	10	.400	10	.900	10	.900	10	0.000	10	1.000	10	.700
Before--Q	A	36	.833	36	.944	36	.639	36	.917	36	.722	36	.889	36	.944	36	.722	36	1.000	36	.833
	NA	1	0.000	1	1.000	1	0.000	1	0.000	1	0.000	1	1.000	1	1.000	1	1.000	1	0.000	1	1.000
After--Q	A	31	.839	31	.968	31	.645	31	.903	31	.774	31	.871	31	.968	31	.774	31	.968	31	1.000
	NA	6	.667	6	.833	6	.500	6	.833	6	.333	6	1.000	6	.833	6	.500	6	1.000	6	0.000

Table 47

Intersubtest Conditional Probabilities for Mastering Tests,

Criterion at 80%

WTRSD Comprehension Level C, Forms P & Q

(N = 40)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability																	
		Detail: Negatives P	N	Para-phrase: Negatives P	N	Topic: Without Organizer P	N	Conclu-sions: One Rela-tionship P	N	Before/ After P	N	Detail: Negatives Q	N	Para-phrase: Negatives Q	N	Topic: Without Organizer Q	N	Conclu-sions: One Rela-tionship Q	N	Before/ After Q
Detail: Negatives--P	A	40	1.000	40	.925	40	.975	40	.825	40	.775	40	1.000	40	.825	40	.975	40	.625	40	.850
	NA	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
Paraphrase: Negatives--P	A	37	1.000	37	1.000	37	.973	37	.811	37	.784	37	1.000	37	.865	37	.973	37	.649	37	.865
	NA	3	1.000	3	0.000	3	1.000	3	1.000	3	.667	3	1.000	3	.333	3	1.000	3	.333	3	.667
Topic: Without Organizer--P	A	30	1.000	39	.923	39	1.000	39	.846	39	.769	39	1.000	39	.821	39	.974	39	.641	39	.872
	NA	1	1.000	1	1.000	1	0.000	1	0.000	1	1.000	1	1.000	1	1.000	1	1.000	1	0.000	1	0.000
Conclusions: One Relation-ship--P	A	33	1.000	33	.909	33	1.000	33	1.000	33	.758	33	1.000	33	.848	33	.970	33	.636	33	.879
	NA	7	1.000	7	1.000	7	.857	7	0.000	7	.857	7	1.000	7	.714	7	1.000	7	.571	7	.714
Before/ After--P	A	31	1.000	31	.935	31	.968	31	.806	31	1.000	31	1.000	31	.839	31	.968	31	.677	31	.935
	NA	9	1.000	9	.889	9	1.000	9	.889	9	0.000	9	1.000	9	.778	9	1.000	9	.444	9	.556
Detail: Negatives--Q	A	40	1.000	40	.925	40	.975	40	.825	40	.775	40	1.000	40	.825	40	.975	40	.625	40	.850
	NA	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0	0.000
Paraphrase: Negatives--Q	A	33	1.000	33	.970	33	.970	33	.848	33	.788	33	1.000	33	1.000	33	1.000	33	.697	33	.879
	NA	7	1.000	7	.714	7	1.000	7	.714	7	.714	7	1.000	7	0.000	7	.857	7	.286	7	.714
Topic: Without Organizer--Q	A	39	1.000	39	.923	39	.974	39	.821	39	.769	39	1.000	39	.846	39	1.000	39	.615	39	.846
	NA	1	1.000	1	1.000	1	1.000	1	1.000	1	1.000	1	1.000	1	0.000	1	0.000	1	1.000	1	1.000
Conclusions: One Relation-ship--Q	A	25	1.000	25	.960	25	1.000	25	.840	25	.840	25	1.000	25	.920	25	.960	25	1.000	25	.920
	NA	15	1.000	15	.867	15	.933	15	.800	15	.667	15	1.000	15	.667	15	1.000	15	0.000	15	.733
Before/ After--Q	A	34	1.000	34	.941	34	1.000	34	.853	34	.853	34	1.000	34	.853	34	.971	34	.676	34	1.000
	NA	6	1.000	6	.833	6	.833	6	.667	6	.333	6	1.000	6	.667	6	1.000	6	.333	6	0.000

Table 48

Intersubtest Conditional Probabilities for Mastering Tests,
Criterion at 80%
WTRSD Comprehension Level D, Forms P & Q
(N = 36)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability																					
		Direct Clues—P		Detail: Passives—P		Para-phrase Passives—P		Topic: Relevant Informa-tion—P		Cause-Effect—P		Calendar Markers—P		Direct Clues—Q		Detail: Passives—Q		Para-phrase Passives—Q		Topic: Relevant Informa-tion—Q		Cause-Effect—Q		Calendar Markers—Q	
		N	P	N	P	N	P	N	P	N	P	N	P	N	Q	N	Q	N	Q	N	Q	N	Q	N	Q
Direct Clues--P	A	24	1.000	24	.958	24	.875	24	.750	24	.583	24	.625	24	.875	24	.958	24	.792	24	.625	24	.667	24	.667
	NA	12	0.000	12	1.000	12	.833	12	.667	12	.333	12	.667	12	.667	12	1.000	12	.667	12	.750	12	.417	12	.750
Detail: Passives--P	A	35	.857	35	1.000	35	.857	35	.714	35	.486	35	.657	35	.800	35	.971	35	.743	35	.686	35	.571	35	.686
	NA	1	1.000	1	0.000	1	1.000	1	1.000	1	1.000	1	0.000	1	1.000	1	1.000	1	1.000	1	0.000	1	1.000	1	1.000
Paraphrase: Passives--P	A	31	.677	31	.968	31	1.000	31	.742	31	.484	31	.613	31	.871	31	1.000	31	.806	31	.677	31	.613	31	.742
	NA	5	.600	5	1.000	5	0.000	5	.600	5	.600	5	.800	5	.400	5	.800	5	.400	5	.600	5	.400	5	.400
Topic: Relevant Information--P	A	26	.692	26	.962	26	.885	26	1.000	26	.538	26	.731	26	.885	26	1.000	26	.769	26	.808	26	.654	26	.769
	NA	10	.600	10	1.000	10	.800	10	0.000	10	.400	10	.400	10	.600	10	.900	10	.700	10	.300	10	.400	10	.500
Cause-Effect--P	A	18	.778	18	.944	18	.833	18	.778	18	1.000	18	.667	18	.833	18	1.000	18	.778	18	.667	18	.722	18	.722
	NA	18	.556	18	1.000	18	.889	18	.667	18	0.000	18	.611	18	.778	18	.944	18	.722	18	.667	18	.444	18	.667
Calendar Markers--P	A	23	.652	23	1.000	23	.826	23	.826	23	.522	23	1.000	23	.826	23	.957	23	.783	23	.739	23	.696	23	.870
	NA	13	.692	13	.923	13	.923	13	.538	13	.462	13	0.000	13	.769	13	1.000	13	.692	13	.538	13	.385	13	.385
Direct Clues--Q	A	29	.724	29	.966	29	.931	29	.793	29	.517	29	.655	29	1.000	29	1.000	29	.828	29	.690	29	.586	29	.759
	NA	7	.429	7	1.000	7	.571	7	.429	7	.429	7	.571	7	0.000	7	.857	7	.429	7	.571	7	.571	7	.429
Detail: Passives--Q	A	35	.857	35	.971	35	.886	35	.743	35	.514	35	.629	35	.829	35	1.000	35	.771	35	.686	35	.600	35	.714
	NA	1	1.000	1	1.000	1	0.000	1	0.000	1	0.000	1	1.000	1	0.000	1	0.000	1	0.000	1	0.000	1	0.000	1	0.000
Paraphrase: Passives--Q	A	27	.704	27	.963	27	.926	27	.741	27	.519	27	.667	27	.889	27	1.000	27	1.000	27	.778	27	.704	27	.741
	NA	9	.556	9	1.000	9	.667	9	.667	9	.444	9	.556	9	.556	9	.889	9	0.000	9	.333	9	.222	9	.556
Topic: Relevant Information--Q	A	24	.625	24	1.000	24	.875	24	.875	24	.500	24	.708	24	.833	24	1.000	24	.875	24	1.000	24	.708	24	.750
	NA	12	.750	12	.917	12	.833	12	.417	12	.500	12	.500	12	.750	12	.917	12	.500	12	0.000	12	.333	12	.583
Cause-Effect--Q	A	21	.762	21	.952	21	.905	21	.810	21	.619	21	.762	21	.810	21	1.000	21	.905	21	.810	21	1.000	21	.762
	NA	15	.933	15	1.000	15	.800	15	.600	15	.333	15	.467	15	.800	15	.933	15	.533	15	.467	15	0.000	15	.600
Calendar Markers--Q	A	25	.640	25	.960	25	.920	25	.800	25	.520	25	.800	25	.880	25	1.000	25	.800	25	.720	25	.640	25	1.000
	NA	11	.727	11	1.000	11	.727	11	.545	11	.455	11	.273	11	.636	11	.909	11	.636	11	.545	11	.455	11	0.000

Table 49

Intersubtest Conditional Probabilities for Mastering Tests,
Criterion at 80%
WTRSD Comprehension Level E, Forms P & Q
(N = 50)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability																					
		Direct Clues: Application		Detail: Clauses		Paraphrase Complex Sentences		Central Thought: With Organizer		Conclusions: Direct Relationships		Explicit Cues		Direct Clues: Application		Detail: Clauses		Paraphrase Complex Sentences		Central Thought: With Organizer		Conclusions: Direct Relationships		Explicit Cues	
		N	P	N	P	N	P	N	P	N	P	N	P	N	Q	N	Q	N	Q	N	Q	N	Q	N	Q
Direct Clues: Application--P	A	32	1.000	32	.938	32	.938	32	.875	32	.781	32	.781	32	.750	32	1.000	32	1.000	32	.969	32	.844	32	.531
	NA	18	0.000	18	.889	18	.889	18	.667	18	.500	18	.333	18	.278	18	.889	18	.778	18	.611	18	.667	18	.389
Detail: Clauses--P	A	46	.652	46	1.000	46	.978	46	.848	46	.739	46	.674	46	.609	46	.978	46	.957	46	.870	46	.783	46	.522
	NA	4	.500	4	0.000	4	.250	4	.250	4	0.000	4	0.000	4	.250	4	.750	4	.500	4	.500	4	.750	4	0.000
Paraphrase: Complex Sentences--P	A	46	.652	46	.978	46	1.000	46	.848	46	.739	46	.674	46	.609	46	.957	46	.957	46	.870	46	.804	46	.522
	NA	4	.500	4	.250	4	0.000	4	.250	4	0.000	4	0.000	4	.250	4	1.000	4	.500	4	.500	4	.500	4	0.000
Central Thought: With Organizer--P	A	40	.700	40	.975	40	.975	40	1.000	40	.800	40	.750	40	.675	40	.975	40	.925	40	.925	40	.825	40	.550
	NA	10	.400	10	.700	10	.700	10	0.000	10	.200	10	.100	10	.200	10	.900	10	.900	10	.500	10	.600	10	.200
Conclusions: Direct Relationships--P	A	34	.735	34	1.000	34	1.000	34	.941	34	1.000	34	.765	34	.735	34	.971	34	1.000	34	.941	34	.853	34	.618
	NA	16	.438	16	.750	16	.750	16	.500	16	0.000	16	.313	16	.250	16	.938	16	.750	16	.625	16	.625	16	.188
Explicit Cues--P	A	31	.806	31	1.000	31	1.000	31	.968	31	.839	31	1.000	31	.742	31	1.000	31	1.000	31	.968	31	.839	31	.581
	NA	19	.368	19	.789	19	.789	19	.526	19	.421	19	0.000	19	.316	19	.895	19	.789	19	.632	19	.684	19	.316
Direct Clues: Application--Q	A	29	.828	29	.966	29	.966	29	.931	29	.862	29	.793	29	1.000	29	1.000	29	1.000	29	1.000	29	.862	29	.655
	NA	21	.381	21	.857	21	.857	21	.619	21	.429	21	.381	21	0.000	21	.905	21	.810	21	.619	21	.667	21	.238
Detail: Clauses--Q	A	48	.667	48	.938	48	.917	48	.813	48	.688	48	.646	48	.604	48	1.000	48	.938	48	.875	48	.792	48	.500
	NA	2	0.000	2	.500	2	1.000	2	.500	2	.500	2	0.000	2	0.000	2	0.000	2	.500	2	0.000	2	.500	2	0.000
Paraphrase: Complex Sentences--Q	A	46	.696	46	.957	46	.957	46	.804	46	.739	46	.674	46	.630	46	.978	46	1.000	46	.891	46	.804	46	.522
	NA	4	0.000	4	.500	4	.500	4	.750	4	0.000	4	0.000	4	0.000	4	.750	4	0.000	4	.250	4	.400	4	0.000
Central Thought: With Organizer--Q	A	42	.738	42	.952	42	.952	42	.881	42	.762	42	.714	42	.690	42	1.000	42	.976	42	1.000	42	.786	42	.548
	NA	8	.125	8	.750	8	.750	8	.375	8	.250	8	.125	8	0.000	8	.750	8	.625	8	0.000	8	.750	8	.125
Conclusions: Direct Relationships--Q	A	39	.692	39	.923	39	.949	39	.846	39	.744	39	.667	39	.641	39	.974	39	.949	39	.846	39	1.000	39	.513
	NA	11	.455	11	.909	11	.818	11	.636	11	.455	11	.455	11	.364	11	.909	11	.818	11	.818	11	0.000	11	.364
Explicit Cues--Q	A	24	.708	24	1.000	24	1.000	24	.917	24	.875	24	.750	24	.792	24	1.000	24	1.000	24	.958	24	.833	24	1.000
	NA	26	.577	26	.846	26	.846	26	.692	26	.500	26	.500	26	.385	26	.923	26	.846	26	.731	26	.731	26	0.000

Table 50

Intersubtest Conditional Probabilities for Mastering Tests,
Criterion at 80%
WTRSD Comprehension Level F, Forms P & Q
(N = 46)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability																	
		Indirect Clues: Application--P		Paraphrase Prepositions--P		Central Thought: Without Organizer--P		Conclusions: Indirect Relationships--P		Implicit Cues--P		Indirect Clues: Application--Q		Paraphrase Prepositions--Q		Central Thought: Without Organizer--Q		Conclusions: Indirect Relationships--Q		Implicit Cues--Q	
		N	P	N	P	N	P	N	P	N	P	N	Q	N	Q	N	Q	N	Q	N	Q
Indirect Clues: Application--P	A	11	1.000	11	.818	11	.727	11	.727	11	.818	11	.455	11	1.000	11	.727	11	.636	11	.091
	NA	35	0.000	35	.829	35	.257	35	.314	35	.171	35	.171	35	.657	35	.571	35	.257	35	.086
Paraphrase: Prepositions--P	A	38	.237	38	1.000	38	.368	38	.421	38	.316	38	.263	38	.763	38	.605	38	.342	38	.105
	NA	8	.250	8	0.000	8	.375	8	.375	8	.375	8	.125	8	.625	8	.625	8	.375	8	0.000
Central Thought: Without Organizer--P	A	17	.471	17	.824	17	1.000	17	.647	17	.412	17	.412	17	.941	17	.706	17	.588	17	.059
	NA	29	.103	29	.828	29	0.000	29	.276	29	.276	29	.138	29	.621	29	.552	29	.207	29	.103
Conclusions: Indirect Relationships--P	A	19	.421	19	.842	19	.579	19	1.000	19	.474	19	.368	19	.842	19	.842	19	.632	19	.105
	NA	27	.111	27	.815	27	.222	27	0.000	27	.222	27	.148	27	.667	27	.444	27	.148	27	.074
Implicit Cues--P	A	15	.600	15	.800	15	.467	15	.600	15	1.000	15	.333	15	.733	15	.667	15	.533	15	.133
	NA	31	.065	31	.839	31	.323	31	.323	31	0.000	31	.194	31	.742	31	.581	31	.258	31	.065
Indirect Clues: Application--Q	A	11	.455	11	.909	11	.636	11	.636	11	.455	11	1.000	11	1.000	11	.727	11	.636	11	0.000
	NA	35	.171	35	.800	35	.286	35	.343	35	.286	35	0.000	35	.657	35	.571	35	.257	35	.114
Paraphrase: Prepositions--Q	A	34	.324	34	.853	34	.471	34	.471	34	.324	34	.324	34	1.000	34	.706	34	.441	34	.088
	NA	12	0.000	12	.750	12	.083	12	.250	12	.333	12	0.000	12	0.000	12	.333	12	.083	12	.083
Central Thought: Without Organizer--Q	A	28	.286	28	.821	28	.429	28	.571	28	.357	28	.286	28	.857	28	1.000	28	.464	28	.071
	NA	18	.167	18	.833	18	.278	18	.167	18	.278	18	.167	18	.556	18	0.000	18	.167	18	.111
Conclusions: Indirect Relationships--Q	A	16	.438	16	.813	16	.625	16	.750	16	.500	16	.438	16	.938	16	.813	16	1.000	16	.188
	NA	30	.133	30	.833	30	.233	30	.233	30	.233	30	.133	30	.633	30	.500	30	0.000	30	.033
Implicit Cues--Q	A	4	.250	4	1.000	4	.250	4	.500	4	.500	4	0.000	4	.750	4	.500	4	.750	4	1.000
	NA	42	.238	42	.810	42	.381	42	.405	42	.310	42	.262	42	.738	42	.619	42	.310	42	0.000

Table 51

Intersubtest Conditional Probabilities for Mastering Tests,
Criterion at 80%
WTRSD Comprehension Level G, Forms P & Q
(N = 50)

A child applying (A) or not applying (NA) one of these tests		Can be expected to master these tests with indicated probability											
		P						Q					
		N	Uncommon Meanings	N	Conclusions: Supported/ Unsupported	N	Implied/ Stated Events	N	Uncommon Meanings	N	Conclusions: Supported/ Unsupported	N	Implied Stated Events
Uncommon Meanings--P	A	44	1.000	44	.523	44	.364	44	.955	44	.500	44	.386
	NA	6	0.000	6	0.000	6	0.000	6	.333	6	0.000	6	.167
Conclusions: Supported/ Unsupported--P	A	23	1.000	23	1.000	23	.609	23	1.000	23	.652	23	.565
	NA	27	.778	27	0.000	27	.074	27	.778	27	.259	27	.185
Implied/ Stated Events--P	A	16	1.000	16	.875	16	1.000	16	1.000	16	.688	16	.563
	NA	34	.824	34	.265	34	0.000	34	.824	34	.324	34	.265
Uncommon Meanings--Q	A	44	.955	44	.523	44	.364	44	1.000	44	.477	44	.409
	NA	6	.333	6	0.000	6	0.000	6	0.000	6	.167	6	0.000
Conclusions: Supported/ Unsupported--Q	A	22	1.000	22	.682	22	.500	22	.955	22	1.000	22	.636
	NA	28	.786	28	.286	28	.179	28	.821	28	0.000	28	.143
Implied/ Stated Events--Q	A	18	.944	18	.722	18	.500	18	1.000	18	.778	18	1.000
	NA	32	.944	32	.313	32	.219	32	.813	32	.250	32	0.000

REFERENCES

Junker, L., & Stewart, D. An index of parallel forms of criterion-referenced tests as a function of the distribution of test scores. Project Paper. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1973.

Klopp, P. M. Preliminary Technical Manual for forms P & Q of the Wisconsin Tests of Reading Skill Development. Working Paper No. 213. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1977.

Otto, W., & Askov, E. The Wisconsin Design for Reading Skill Development: Rationale and Guidelines. Minneapolis: National Computer Systems, 1974.

Otto, W., & Kamm, K. The Wisconsin Design for Reading Skill Development: Teacher's Planning Guide--Comprehension. Minneapolis: National Computer Systems, 1977.

Stewart, D. Technical Manual for the Wisconsin Tests of Reading Skill Development: Word Attack and Study Skills. Working Paper No. 190. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

APPENDIX A

STATEMENT OF SKILLS AND OBJECTIVES FOR THE WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT: COMPREHENSION COMMERCIAL EDITION

Level A

1. Derives meaning from passages
   a. Identifies a topic: pictures (1)

      Objective: The child identifies a picture which best illustrates the topic of an orally presented passage that has an organizer (topic sentence) and only relevant information (details that support or describe the topic).

   b. Identifies relationships (2)

      Objective: The child identifies a picture that illustrates an orally presented statement of a class inclusive relationship (e.g., All the boats are blue) or a comparison/contrast relationship (e.g., The red book is bigger than the blue book).

2. Determines sequence: first or last event (3)

   Objective: Given an orally presented selection with clearly defined sequential events, the child determines that one event occurs first and another event occurs last.

Level B

1. Derives meaning from sentences: notes detail (1)

   Objective: Given a short written-oral selection of simple sentences, the child identifies details: verbs, modifiers, direct objects, and objects of positional prepositions (e.g., In the sentence, "John has a red wagon," the child determines the color of the wagon).

2. Derives meaning from passages
   a. Identifies a topic: with organizer (2)

      Objective: The child identifies the topic of a written-oral passage that has an organizer (topic sentence) and only relevant information (details that support or describe the topic).

   b. Predicts outcomes (3)

      Objective: The child synthesizes two or more events to identify a logical outcome.

3. Determines sequence
   a. Identifies event: before (4)

      Objective: Given an orally presented selection with clearly defined sequential events, the child determines that one event occurred before another event.

   b. Identifies event: after (5)

      Objective: Given an orally presented selection with clearly defined sequential events, the child determines that one event occurred after another event.

Level C

1. Derives meaning from sentences

   a. Notes detail in positive and negative sentences (1)

      Objective: Given a short, written selection of simple positive and negative sentences, the child identifies details: pronouns, indirect objects, objects of infinitives, and objects of directional prepositions (e.g., In the sentence, "John gave the book to Jim," the child determines who received the book).

   b. Paraphrases positive and negative sentences (2)

      Objective: The child restates (1) simple positive sentences in the positive and in the negative; and (2) simple negative sentences in the positive and in the negative by rearranging the order of words in the sentence and/or by substituting for one or more words.

2. Derives meaning from passages

   a. Identifies a topic: without organizer (3)

      Objective: The child identifies the topic of a passage with no organizer (topic sentence) and only relevant information (details that support or describe the topic).

   b. Identifies conclusions: one relationship (4)

      Objective: The child determines deductively whether a conclusion is correct or incorrect from the relationship and instance directly stated in a passage.

3. Determines sequence: event before or after (5)

   Objective: Given a written passage with clearly defined sequential events, the child determines that one event occurred before or after another event.

Level D

1. Determines word meaning: identifies direct context clues (1)

   Objective: The child uses explicit context clues (e.g., synonym, definition, equivalent phrase, summary) to determine the meaning of an unfamiliar word in context.

2. Derives meaning from sentences

   a. Notes detail in active and passive voice sentences (2)

      Objective: The child attends to detail in sentences which are written in the active or passive voice and contain (1) temporal and/or abstract prepositional phrases and (2) either a terminal clause or an embedded part.

   b. Paraphrases active and passive voice sentences (3)

      Objective: The child restates sentences which are written in the active and passive voice and contain no more than one subordinate clause by rearranging the order of words in the sentence and/or by substituting for one or more words.

3. Derives meaning from passages

   a. Identifies relevant information (4)

      Objective: The child determines whether information is relevant or irrelevant to the topic of a passage that has an organizer (topic sentence) and both relevant and irrelevant information.

b. Identifies cause-effect relationships (5)

Objective: The child determines whether a given relationship is the cause-effect relationship stated in a passage.

4. Determines sequence: calendar markers (6)

Objective: Given a written passage, the child determines the order of events when the order is indicated by calendar markers (e.g., spring, afternoon).

Level E

1. Determines word meaning

a. Identifies word parts: prefixes (1)

Objective: The child recognizes that a prefix is a meaning-bearing unit that can modify the meaning of a word by signifying one of the following: range, time, number, approval or support, disapproval or opposition.

b. Identifies direct context clues: application (2)

Objective: The child determines the meaning of an unfamiliar word in a context which contains explicit clues (e.g., synonym, definition, equivalent phrase, summary) and then applies that meaning in a second context where the same unfamiliar word appears but no explicit clues are present.

2. Derives meaning from sentences

a. Notes detail in sentences with more than one subordinate clause (3)

Objective: The child attends to detail in sentences which contain no more than two subordinate clauses.

b. Paraphrases complex sentences (4)

Objective: The child restates sentences which contain no more than two subordinate clauses by rearranging and/or by substituting for more than a short phrase.

3. Derives meaning from passages

a. Identifies central thought: with organizer (5)

Objective: The child identifies a central thought of a passage that contains an organizer (topic sentence) and both relevant and irrelevant information.

b. Identifies conclusions: direct relationships (6)

Objective: The child identifies relationships that are directly stated in a passage and then determines whether given conclusions based on those relationships are right or wrong. (A relationship that is directly stated is one that is given in a single statement.)

4. Determines sequence: explicit cues (7)

Objective: Given a written passage, the child determines the order of specific events when the order is indicated by explicit sequential markers (e.g., next, finally).

Level F

1. Determines word meaning

   a. Identifies word parts: suffixes (1)

      Objective: The child recognizes that a suffix may serve (1) to modify the meaning of a base word and/or (2) to identify the grammatical function of the base word.

   b. Identifies indirect context clues: application (2)

      Objective: The child determines the meaning of an unfamiliar word in a context which contains implicit clues (e.g., cause-effect, contrast, example, modifying phrase) and then applies that meaning in a second context where the same unfamiliar word appears but no implicit clues are present.

2. Derives meaning from sentences: paraphrases complex sentences with two or more prepositional phrases (3)

   Objective: The child paraphrases complex sentences containing two or more prepositional phrases by rearranging words and by substituting for more than a short phrase.

3. Derives meaning from passages

   a. Identifies central thought: without organizer (4)

      Objective: The child identifies a central thought of a passage that contains no organizer (topic sentence) and has both relevant and irrelevant information.

   b. Identifies conclusions: indirect relationships (5)

      Objective: The child identifies relationships that are indirectly stated in a passage and then determines whether given conclusions based on those relationships are right or wrong. (A relationship that is indirectly stated is one that requires synthesis of information from more than one statement.)

4. Determines sequence: implicit cues (6)

   Objective: The child determines where a specific event occurs within the framework of a series of explicit events by attending to implicit sequential cues.

Level G

1. Determines word meaning

   a. Identifies word parts: combining forms (1)

      Objective: The child recognizes that combining forms (e.g., bio, graph) are meaning-bearing units that join together to constitute words.

   b. Identifies word parts: word roots (2*) (optional)

      Objective: The child recognizes that some words consist of prefixes or suffixes combined with word roots (e.g., fac, sta).

*Note: If word roots are included in the developmental reading program, this skill can be taught; if word roots are not included, this skill can be omitted. Although the ability to recognize word roots probably has little inherent value, the child who has become familiar with them may have a broader base for vocabulary development.

c. Identifies context clues: obscure meanings (3)

Objective: The child uses direct and indirect context clues to determine the obscure meaning of a familiar word in context.

2. Derives meaning from sentences: rewrites sentences (4)

Objective: The child paraphrases one long sentence by rewriting it into four or more sentences and also paraphrases several shorter sentences by rewriting them into one long sentence.

3. Derives meaning from passages

a. Generates central thought (5)

Objective: The child generates a central thought of a passage that may or may not have an organizer (topic sentence) and has both relevant and irrelevant information.

b. Recognizes supported and unsupported conclusions (6)

Objective: The child identifies relationships that are either directly or indirectly stated in a passage and then determines whether given conclusions based on those relationships are supported or unsupported. (A relationship that is directly stated is one that is given in a single statement. A relationship that is indirectly stated is one that requires synthesis of information from more than one statement.)

c. Modifies conclusions (7)

Objective: Given passages that include one or more directly or indirectly stated relationships, the child forms a tentative conclusion and then modifies that conclusion as additional facts or evidence is presented.

4. Determines sequence: implied and stated events (8)

Objective: The child determines where a specific event occurs within the framework of a series of events by attending to implicit or explicit sequential cues.

APPENDIX B

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q

ESTIMATED SUMMARY STATISTICS FOR FORMS P and Q

Tryout Population

The large-scale test tryouts, step 4 in the R & D Center test development model, were conducted in schools that were using the developmental edition of the Design Comprehension element and in schools that were not. The proportions from each site type were equal; 50 percent of the testees were Design Comprehension-experienced (75 students) and 50 percent were Design Comprehension-naive (75 students), for a total tryout population of approximately 150 for each test.

The other tryout variable was grade level. Each test was tried out at the approximate grade level for its intended use. Some grade level adjustments were necessary, however, because the tryouts were conducted throughout the 1975-76 school year, generally in skill level order. For example, although Level A is intended for the end of kindergarten, the Level A tryouts were conducted in Fall and mid-Winter with a first grade rather than a beginning kindergarten population. The Level B tests were tried out at grade 2, Level C at grade 3, and Level D at grade 4. Most of the Level E, F, and G tests were tried out in Spring 1976 at grades 4, 5, and 6, respectively. A few Level E and F tests were tried out in Fall 1975 or Winter 1976 at grades 5 and 6, respectively. The Level G Sequence test was tried out at both grades 6 and 7.

The Design Comprehension-experienced testees attended six schools, five in Wisconsin and one in suburban Denver, Colorado. All of these

schools had participated in the 1973-74 Design Comprehension small-scale field test. Based upon their standardized test results obtained during the field test, the six schools were categorized as low average (two schools), average (two schools), or nigh average (two schools) in Comprehension reading achievement. Each tryout was conducted with testees from each of the three achievement categories. Each school, however, did not participate in every tryout. Within each achievement category, the tryout candidates were students at alternate grade levels in each school. For example, in the high average category, candidates came from grades 1, 3, or 5 in one school and from grades 2, 4, or 6 in the other school. The students who participated in each tryout were randomly selected from school-provided class lists.

Students from four schools near Madison constituted the Design Comprehension-naive test group. Typical reading achievement in these schools was judged to be average to above average. A high enrollment at each grade level permitted sampling, so that any one student participated in approximately half, instead of all, the tryouts at that grade level. The district curriculum coordinator was familiar with the overall sampling procedure and selected the students for each tryout.

Tryouts for the seven completion tests--E1, F1, G1, G2, G4, G5, and G7--were also conducted in both site types. One tryout difference, however, was that the testees in the Design Comprehension-experienced schools were teacher-selected. Another difference was population size, much smaller here than for the multiple-choice tests because of the scoring usability information needed (see section II). Each test form for Tests E1, F1, G1, G2, G4, and G5 was tried out with 12-24 students. Test G7, which was

individually administered, was tried out with 8 students.

Summary Statistics

Estimated summary statistics for the 33 multiple-choice tests are presented by skill level in Tables A to G. No data are reported for the other seven tests, since they were analyzed in a different manner from the multiple-choice tests, and since several revisions were made based upon the tryout results.

The data for several tests in Tables A to G require amplification. First, Tests B3, C4, F6, and G3 were retained from the developmental edition; the grade levels for the answer documents from which the data were obtained are unknown. Since only Form P (Form I, developmental edition) was machine-scored, no comparable data are available for Form Q (Form II, developmental edition). Second, several tests were re-leveled, based upon the tryout results. Because very high means were obtained in the tryout of the Levels A and B Relationships tests, items from the two tests were combined to assemble Test A2 (Table A). Since the final assembled test was not tried out, the data for Test A2 represent two different populations and grade levels.

Tests C4 and C5 (Table C) were originally Level D tests. Data obtained during the 1973-74 small-scale field test suggested that Test 4, Conclusions: One Relationship, could be re-leveled down to Level C. Data presented for Test 5, Sequence--Before/After, were obtained at grade 4; the data for Tests C1, C2, and C3 were all obtained at grade 3. Tests D5 and D6 (Table D), on the other hand, were originally Level C tests. Data reported for these two tests were obtained at grade 3; data for the other

Level D tests were all obtained at grade 4. The last explanation refers to Test G8. Although this Sequence test was tried out at both grades 6 and 7, the quantities of missing data on grade 6 answer documents necessitated using grade 7 data for item selection and test assembly.

TABLE A

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL A

Test	No. Items Per Form	Form	Mean Raw Score	Mean Percent Correct	Std. Dev.	r
1--Central Thought--Topic: Pictures	12	P	11.16	93.0	1.61	.78
		Q	11.23	93.6	1.61	.79
2--Relationships	15	P	14.78	98.5	.52	.23
		Q	14.68	97.9	.62	.21
3--Sequence--First/Last Event	15	P	14.14	94.3	2.21	.89
		Q	14.17	94.5	2.27	.91

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

TABLE B

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL B

Test	No. Items Per Form	Form	Mean Raw Score	Mean Percent Correct	Std. Dev.	r
1--Detail	15	P	12.79	85.3	2.38	.75
		Q	12.84	85.6	2.52	.77
2--Central Thought--Topic: With Organizer	15	P	13.05	87.0	2.58	.81
		Q	13.06	87.1	2.53	.80
3--Relationships/Conclusions: Outcomes[a]	12	P	7.04	58.7	2.91	.73
		Q				
4--Sequence--Before	20	P	17.63	88.1	3.08	.83
		Q	17.56	87.8	3.11	.83
5--Sequence--After	20	P	17.19	85.9	3.28	.82
		Q	17.14	85.7	3.30	.82

[a] Test was retained from the 1974-76 Developmental Edition. Data source: all Level B answer documents, Developmental Edition, Form I only, scored by National Computer Systems between 9/74 and 1/75.

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

TABLE C

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL C

Test	No. Items Per Form	Form	Mean Raw Score	Mean Percent Correct	Std. Dev.	r
1--Detail: Negatives	15	P	13.55	90.3	2.33	.82
		Q	13.63	90.9	2.21	.81
2--Paraphrase: Negatives	20	P	16.37	81.8	3.12	.75
		Q	16.37	81.8	3.09	.74
3--Central Thought--Topic: Without Organizer	15	P	13.13	87.5	3.05	.89
		Q	13.12	87.5	3.01	.88
4--Conclusions: One Relationship[a]	20	P	18.01	90.0	2.57	.77
		Q				
5--Sequence--Before/After	20	P	17.22	86.1	2.96	.77
		Q	17.20	86.0	2.94	.77

[a] Test was retained from the 1974-76 Developmental Edition. Data source: all Level D answer documents, Developmental Edition, Form I only, scored by National Computer Systems between 9/74 and 1/75.

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

TABLE D

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL D

Test	No. Items Per Form	Form	Mean Raw Score	Mean Percent Correct	Std. Dev.	r
1--Context--Direct Clues	17	P	13.64	80.2	3.22	.80
		Q	13.70	80.6	3.29	.81
2--Detail: Passives	20	P	18.51	92.5	2.68	.85
		Q	18.49	92.4	2.85	.87
3--Paraphrase: Passives	20	P	16.40	82.0	3.15	.75
		Q	16.45	82.2	3.11	.75
4--Central Thought--Topic: Relevant Information	20	P	16.26	81.3	3.30	.77
		Q	16.29	81.4	3.40	.79
5--Relationships/Conclusions: Cause-Effect	20	P	13.43	67.1	3.74	.77
		Q	13.41	67.0	3.77	.77
6--Sequence--Calendar Markers	20	P	15.41	77.0	3.54	.77
		Q	15.34	76.7	3.52	.76

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

TABLE E

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL E

Test	No. Items Per Form	Form	Mean Raw Score	Mean Percent Correct	Std. Dev.	r
2--Context--Direct Clues: Application	14	P	10.78	77.0	3.09	.82
		Q	10.78	77.0	3.20	.83
3--Detail: Clauses	20	P	18.14	90.7	3.12	.87
		Q	18.14	90.7	3.19	.88
4--Paraphrase: Complex Sentences	20	P	17.36	86.8	2.89	.77
		Q	17.36	86.8	2.90	.77
5--Central Thought: With Organizer	15	P	12.55	83.7	2.72	.79
		Q	12.58	83.9	2.73	.79
6--Conclusions: Direct Relationships	20	P	15.75	78.7	3.46	.77
		Q	15.81	79.0	3.59	.79
7--Sequence--Explicit Cues	15	P	11.13	74.2	3.63	.84
		Q	10.73	71.5	3.97	.87

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

Data collected for Test 1 were analyzed differently from the other tests at this level and therefore are not listed here.

TABLE F

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, LEVEL F

Test	No. Items Per Form	Form	Mean Raw Score	Percent Correct	Std. Dev.	r
2--Context--Indirect Clues: Application	14	P	9.69	69.2	2.83	.70
		Q	9.69	69.2	2.81	.69
3--Paraphrase: Prepositions	20	P	17.42	87.1	2.33	.64
		Q	17.46	87.3	2.29	.64
4--Central Thought: Without Organizer	14	P	12.12	86.6	2.14	.72
		Q	12.20	87.1	2.11	.72
5--Conclusions: Indirect Relationships	20	P	15.25	76.2	3.30	.72
		Q	15.32	76.6	3.20	.70
6--Sequence--Implicit Cues[a]	12	P	8.35	69.6	3.06	.81
		Q				

[a] Test was retained from the 1974-76 Developmental Edition. Data source: all Level F answer documents, Developmental Edition, Form I only, scored by National Computer Systems between 9/74 and 1/75.

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

Data collected for Test 1 were analyzed differently from the other tests at this level and therefore are not listed here.

TABLE G

ESTIMATED SUMMARY STATISTICS FOR FORMS P AND Q
1977 COMMERCIAL EDITION OF WTRSD: COMPREHENSION, Level G

Test	No. Items Per Form	Form	Mean Raw Score	Percent Correct	Std. Dev.	r
3--Context--Uncommon Meanings[a]	12	P	10.59	88.2	2.18	.81
		Q				
6--Conclusions: Supported/ Unsupported	21	P	15.41	73.4	3.24	.68
		Q	15.23	72.5	3.15	.68
8--Sequence--Implied/ Stated Events	15	P	7.87	52.5	3.91	.82
		Q	7.82	52.1	3.76	.80

[a] Test was retained from the 1974-76 Developmental Edition. Data source: all Level G answer documents, Developmental Edition, Form I only, scored by National Computer Systems between 9/74 and 1/75.

NOTE: Forms P and Q were assembled from the total number of items selected from the tryout data. Student N for the tryout data was approximately 150. Unless footnoted otherwise, the mean, standard deviation, and reliability for each form were estimated from item statistics obtained in tryout.

Data collected for Tests 1, 2, 4, 5, and 7 were analyzed differently from the other tests at this level and therefore are not listed here.

APPENDIX C

ITEM WRITING GUIDELINES FOR THE WISCONSIN TESTS OF READING SKILL DEVELOPMENT: COMPREHENSION COMMERCIAL EDITION

April 1975

Tentative item writing guidelines have been developed for the Comprehension element. These guidelines include four basic aspects: (1) general sentence structures; (2) phrase limitations; (3) sentence length; and (4) vocabulary restrictions. The general sentence structures were developed as part of the Sentence Meaning strand and are explained more fully in the Sentence Meaning Rationale. The guidelines for phrase limitations were developed primarily as a result of information gleaned from studies relating to Sentence Meaning research. The sentence length specifications, including maximum and average sentence length and number of words per passage, are very similar to those used for the Developmental Edition, i.e., the estimates are based on the Spache and Dale-Chall readability formulas. Vocabulary restrictions are also similar to those used for the Developmental Edition. Each of these four aspects is explained in more detail below. The guidelines will be applied at all levels in each strand except where there is a conflict; e.g., the Level C Sequence skill is tested with introductory subordinate clauses which the guidelines don't specify until Level E. "Concept load" will be attended to via consideration of these specifications and the judgment of project and evaluation staff.

1. General Sentence Structures

Level B -- Simple sentences (written-oral)

Level C -- Simple sentences

Level D -- Compound or complex (1 middle or terminal clause) sentences

Level E -- Compound and/or complex (2 clauses) sentences

Level F -- Two clauses with 2+ prepositional phrases

Level G -- One or more sentences of any type

2. Phrase Limitations for Sentences

Level B -- 1 prepositional phrase per sentence
1 infinitive (phrase) per sentence
no gerund/participial phrases
no phrases to left of subject

Maximum of 2 of these structures per sentence

Level C -- Same specifications as for Level B

NOTE: For Levels D-G phrase specifications include the main as well as subordinate clause.

Level D -- 2 prepositional phrases per sentence
2 infinitives (phrases) per sentence
1 gerund/participial phrase per sentence
1 short phrase to left of subject (prepositional)

Maximum of 3 of these structures per sentence

Level E -- 2 prepositional phrases per sentence
2 infinitive (phrases) per sentence
1 gerund/participial phrase per sentence
any phrase to left of subject

Maximum of 4 of these structures per sentence

Level F -- 2 infinitives (phrases) per sentence
1 gerund/participial phrase per sentence

Maximum 4 of these structures per sentence

Level G -- any type of structure appropriate for 6th grade

3. Sentence Length

	Maximum words per passage	Maximum sentence length	Maximum average sentence length
Level B	50	15	9
Level C	50	15	9
Level D	96	20	12
Level E	120	22	15
Level F	150	22	15
Level G	80	27	20

4. Vocabulary Restrictions

There are four categories of words. Any word falling into the first three is considered "familiar." Any word in category IV must be counted as "unfamiliar." The maximum number of unfamiliar words per passage for any level is 10%

Category I. Any word on the Word List* is acceptable.

Category II. Any word that is an appropriate concept and attackable according to the Word Attack skills at the previous level (except for Level B):

	Comprehension	Word Attack
Written/oral	Level B	Level B
	Level C	Level C
	Level D-G	Level D

*The source for the Word List is Carroll, J. B., Davies, P., & Richman, B. American Heritage Word Frequency Book. Boston: Houghton Mifflin, 1971. The first 2,000 most frequent words are designated by "39" at grade level 3 and may be used for Levels A-D. For Levels E, F, and G, grade levels 4, 5, and 6 should be consulted.

Category III. Any word that is very familiar (e.g., everyday usage) to children; yet unattackable (with previous level Word Attack skills), e.g., television.

This category is, admittedly, a "gray area," and perhaps we won't be able to make the proper judgments. But we definitely need, especially at Levels C and D, some latitude beyond the word list and the Level B-C skills for writing the items. One suggestion is to consider allowing only one to three of these words in any one passage. That is, if a given passage has the maximum 10 percent unfamiliar words (Category IV) in it, then only one or two Category III words would be allowed. If the number of unfamiliar words were not maximum, then perhaps the story could contain as many as three or four of Category III words.

Category IV. Any word which is an appropriate concept and unattackable (with previous level Word Attack skills).

APPENDIX D

LIST OF SELECTED DESIGN-RELATED STUDIES

Askov, E. N. Assessment of a system for individualizing reading instruction. Technical Report No. 117. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Chester, R., Hubbard, W. D., & Van Blaricom, G. Field test report of the Self-Directed, Interpretive, and Creative Reading elements of the Wisconsin Design for Reading Skill Development. Technical Report No. 330. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Duffy, G. Maintaining a balance in objective-based reading instruction. In The Reading Teacher, February 1978, 31, 519-523. D

Johnson, S. First-year field test of the Wisconsin Design for Reading Skill Development. Technical Report No. 5, Atlanta, Georgia: Southeastern Educational Laboratory, December 1971.

Kamm, K. The Wisconsin Design for Reading Skill Development: Study Skills--A case study. Technical Report No. 422, October 1977. E

Kamm, K., White, S., & Morrison, B. A report of the procedures used in the implementation of an objective-based reading program in 16 schools. Working Paper No. 246. Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, 1977.

Kling, M. An analysis and evaluation of three management support systems in reading: Fountain Valley, Prescriptive Reading Inventory, and Wisconsin Design for Reading Skill Development. Paper presented at the College Reading Association, Washington, D. C., November 1973.

Klopp, P. The Wisconsin Design for Reading Skill Development: Comprehension--A report on the 1973-74 small-scale field test. Technical Report No. 373. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Klumb, R. W. The effects of monitoring pupil performance and two incentive treatments for teachers on pupils' reading skill development and teachers' attitudes. Technical Report No. 268. Madison: Wisconsin Research and Development Center for Cognitive Learning, August 1973.

Kurth, R. J. Evaluation of an objective-based curriculum in word attack. Technical Report No. 289. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Otto, W., & Askov, E. The Wisconsin Design for Reading Skill Development: Rationale and guidelines. Minneapolis: National Computer Systems, 1974.

Otto, W., & Chester, R. Objective-Based Reading. Reading, Massachusetts: Addison-Wesley, 1976.

Otto, W., Chester, R., McNeil, J., & Meyers, S. Focused Reading Instruction. Reading, Massachusetts: Addison-Wesley, 1974.

Quilling, M. R. The first field test of the Word Attack element of the Wisconsin Design for Reading Skill Development. Technical Memo No. QV-10-70, 1971. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1971.

Quilling, M. & Otto, W. Evaluation of an objective-based curriculum in reading. Journal of Educational Research, 65, 1971, 15-18.

Quilling, M. R., & Wojtal, P. The Wisconsin Design for Reading Skill Development: Study Skills--A report on the pilot test: 1970-72. Working Paper No. 97. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

Rosenkranz, C. I. Comprehension: A reading improvement project for EMRs. Bureau Memorandum XXIV-4. Gillett, Wisconsin: Cooperative Educational Service Agency No. 3, 1973, pp. 25-26.

Rude, R. T. Implementing and field-testing of the Design. In Green, F. P. (Ed.), Investigations Relating to Mature Reading, Twenty-first yearbook of the National Reading Conference. Milwaukee: National Reading Conference, Inc., 1972, 265-272.

Rude, R. T. Sex, intelligence and school reading curriculum as factors influencing summer retention of overall reading ability and specific reading skills of first grade subjects. Technical Report No. 263. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1973.

Sals, D., & Hubbard, W. D. The Wisconsin Design for Reading Skill Development: Study Skills--A report on the Type I field test, 1971-73. Technical Report No. 341. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Stewart, D. Technical manual for the Wisconsin Tests of Reading Skill Development: Word Attack and Study Skills. Working Paper No. 190. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Wisconsin Research and Development Center for Cognitive Learning. Wisconsin Design Makes a Difference. Multiunit Newsletter (February 1972).

Wisconsin Research and Development Center for Cognitive Learning. Field Test Report, 1971-72. Multiunit Newsletter, 1973, 4 (2), 2-4.

Wisconsin Research and Development Center for Cognitive Learning. Evaluation--Wisconsin Design for Reading Skill Development: Word Attack type II field test, 9/70-5/72. January 1974a.

Wisconsin Research and Development Center for Cognitive Learning. Evaluation--Wisconsin Design for Reading Skill Development: Study Skills type II field test, 9/71-5/73. October 1974b.